Learning of Long-Run Risk in International Markets

Steven Wei Ho

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration (Finance) in the Department of Finance of the Kenan-Flagler Business School.

Chapel Hill
2013

Approved by:

Mariano Croce

Riccardo Colacito

Christian Lundblad

David Ravenscraft

Toan Phan

Abstract

STEVEN WEI HO: Learning of Long-Run Risk in International Markets.
(Under the direction of Mariano Croce and Riccardo Colacito.)

I develop a general equilibrium model in which agents from two countries do not observe directly the long-run growth prospects of their economies. Instead, the agents rationally learn the hidden components through the Kalman filter applied to international consumption data. Learning endogenously produces: (i) a rational explanation of international contagion phenomenon, defined as changes in one country's asset prices in response to foreign news, that occurs in the absence of domestic news, (ii) large and time-varying international equity risk premia, and (iii) a resolution of the forward premium anomaly, defined as the tendency of high interest rate currencies to appreciate.

Dedication

I dedicate my dissertation to my family. My name He Wei, written in Chinese characters, 贺威, was given by my late Grandfather, Prof. He Renlin 贺仁麟, who came from Jiangxi Province and worked at Shanghai Normal University 上海师范大学. My late Grandmother, Hao Minzhuang 郝敏莊, who came from Kaifeng 赫舍里氏 of both Manchurian and Mongolian descent, has instilled me the importance of education since I was young. She has always supported me studying abroad despite missing me very much. She recently passed away in November, 2012. To my Father, He Gang 贺钢 of Shanghai Nanlin Normal School 南林师范学校, and to my Mother, Dr. Su Meifang 苏美芳 of Shanghai Huashan Hospital, who made my journey possible. To my maternal Grandmother Diao Suzhen 刁素贞, and my maternal Grandfather Su Yinbao 苏寅宝 who, as an architectural engineer from 绍兴, has served the Shanghai Municipal Government and worked on bridge constructions.

Thanks to my friends and our faith. I had the unusual fortune.

Acknowledgments

I am grateful for all the help I have received. I thank my advisors, Mariano Croce and Riccardo Colacito, for the mentorship that has made my dissertation possible. I am deeply indebted to David Ravenscraft. I also thank my committee members, Christian Lundblad and Toan Phan for providing me with valuable guidance at many occasions. Special thanks to Alexander Arapoglou and Gregory Brown for the teaching opportunity. I am also grateful for the help from David Dicks, Pab Jotikasthira, Robert Connolly and Paolo Fulghieri. I am also indebted to Wayne Landsman and Jennifer Conrad for supporting the doctoral program at Kenan-Flagler Business School, UNC Chapel Hill.

Preface

With this dissertation, I will go back to China in 2013.

I left Shanghai and came to Canada in 1999.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Do agents know the long-run prospect of the economy? I address this question by investigating the international asset pricing implications of an economy in which agents have to learn the long-run prospect of the economy from international historical growth data. To quote from Hansen, Heaton, and Li (2008), "many of the statistical challenges that plague econometricians presumably also plague market participants. Naive application of rational expectations equilibrium concepts may endow investors with too much knowledge about future growth prospects".

A growing body of the literature has documented that long-run growth prospects influence both domestic (Bansal and Yaron, 2004) and international asset prices (Colacito and Croce, 2011). All these models are built around the assumption that consumption is exposed to two sources of shocks perfectly observable to the agents. Specifically, short-run risk is modeled as an i.i.d shock that only affects the economy for one period, whereas long-run risk is modeled as a small, but highly persistent autoregressive process, which is thus expected to affect long-run growth rates. A key question within the long-run risk paradigm is the ability of agents to perfectly detect such a small long-run component.

Earlier papers have investigated this question in a one-country setting (Ai, 2010; Bansal and Shaliastovich, 2011; Croce, Lettau, and Ludvigson, 2012; Johannes, Lochstoer, and Mou, 2010). In this article, I study a two-country economy, each populated by one consumer. The dynamics of consumption growth within each country feature the aforementioned long-run risks. The agents do not know the actual value of each country's specific long-run risk, in the sense that they cannot can accurately break down consumption news into a component which is going to affect the economy for many periods, and a component which represent short-run i.i.d shocks that is only going to affect the consumption stream for one period. The long-run components of the two countries are correlated (Colacito and Croce, 2011), thus innovation in one country can provide information regarding the long-run component of the other country.

The Bayesian learning mechanism employed in this article is the Kalman filter (Kalman and Bucy, 1961; Kalman, 1960). Due to the recursive nature of the Kalman filter, it can also be viewed as the simplest dynamic Bayesian Network (Haykin, 2001; Murphy, 2002). An intuitive interpretation of the Kalman filter approach is that in each period the agent forms a prior regarding the joint probability distribution of the two latent variables, which are the home and foreign long-run risks, and come up with a prediction for the following period; as time goes by these predictions are compared with the realized observations. The agent adjusts the estimates of the latent variables and turns her posterior distribution into a new prior to be used in the subsequent period. Under such a setup, agent from one country would recursively Bayesian update her estimate of her country-specific long-run component by utilizing all available historical information in consumption growth in both countries, and optimizes her asset pricing decisions accordingly. As an extension of the setup, I also investigate the case of learning from both consumption and dividend stream.

I find that learning can account for prominent features of domestic and international financial markets. Specifically, the learning model presented in this article offers a solution of the puzzling contagion phenomenon in the context of a dynamic stochastic equilibrium framework with complete markets and no arbitrage. Contagion is broadly defined as the propagation of shocks in excess of what can be explained by fundamentals (Forbes and Rigobon, 2000). There is no universally accepted explanation for why contagion occurs in equilibrium (Bekaert and Harvey, 2003; Karolyi and Stulz, 1996). This phenomenon is puzzling as one could potentially arbitrage when asset prices deviate from fundamentals. Current explanations offered include investor herding behavior (Cipriani and Guarino, 2008), and the spread of fear or "market sentiments" (Fromlet, 2001). However, such psychological arguments do not prevent an agent, who is perfectly rational and unmoved by sentiments, from undertaking arbitrage. I propose a rational expectations explanation for why contagion may occur in equilibrium.

I define the true long-run components of each country as the economic fundamentals, which are unobservable and thus have to be learned, taking the history of the consumption stream of both countries as observables. The agents of the two countries are endowed with the same information. When agents learn using Kalman filter, they form an estimate of the entire joint distribution of the two latent variables characterized by the means and the variance-covariance matrix. As a result, innovations in consumption streams of both countries, are two sources of information that will be incorporated into the estimate of the long-run persistent component of either one country. Since the growth rates of consumption and their long-run components are correlated across countries, the estimate of home's economic fundamental will be revised in response to foreign innovations, even in the absence of any new piece of information in the home country. Equivalently, Bayesian learning gives rise to contagion as one country may revise its long-run growth expectations solely in response to news coming from abroad. Note that at each point in time, equilibrium prices are determined using agents' estimates of the distribution of fundamentals given all historical information up to that point. This means that an econometrician who observes the entire data series may observe temporary mis-pricing, conditional on her information set which is larger than the agents'.

The model is also able to provide a rational explanation for the forward premium anomaly, i.e. the well-documented tendency of high interest rate currencies to appreciate (Fama, 1984), which is at odds with the prediction of the uncovered interest rate parity relationship. The model in this article can provide a solution due to its ability to endogenously generate time-varying volatility through learning. Take for example the case in which the home country agent believes that the estimate of her long-run component is very unreliable, i.e. it features a large variance. This variance of the estimation error will steadily decrease through learning in the transition toward the steady-state Kalman filter value. During this process, the consumption profile of the home agent becomes less uncertain and thus the interest rate in this country gets relatively larger. By no arbitrage, the currency of the home country become more valuable, being associated to a safer consumption profile, and it is thus expected to appreciate, despite its higher interest rate.

Furthermore, I show that within each country, learning increases the equity risk premium by as much as 22%. This finding confirms and extends the analysis of Croce, Lettau, and Ludvigson (2012) to the case in which investors exploit the cross-sectional dimension of countries in addition to the domestic time-series of consumption. Additionally, when the Kalman filter is off steady-state, the model endogenously generates time-varying uncertainty, as the recursive application of the Bayesian updating changes the accuracy of the estimated long-run risks through time. This means that the model features a time-varying risk-premium.

The rest of the article is organized as follows. Chapter 2 reviews the related literature. In Chapter 3, I describe the model setup, the information structure involved, and the learning. In Chapter 4, I present detailed derivations. In Chapter 5, I present the analytical asset pricing solutions of learning under different information structures. In Chapter 6, I provide both the theoretical arguments and numerical simulation results which would explain contagion in international markets and the forward-premium puzzle. Chapter 7 concludes the article.

Chapter 2

Background Literature

In the Bansal and Yaron (2004) model, there is a small but highly persistent predictable long-run component in consumption growth, which is subject to long-run shocks. They are called long-run shocks since due to the highly persistent nature of the long-run component, even small innovations in this component will induce large cash flow movements in a long time horizon. Thus, the low-frequency movements in consumption growth rates are called the long-run risk (Bansal, Kiku, and Yaron, 2010).

A notable ingredient of the long-run risk literature, which I also employ here, is the Epstein-Zin-Weil preference (Epstein and Zin, 1989; Weil, 1989). Under this recursive preference, agents in each period would optimize the tradeoff between utility of current period and continuation utility derived from all future periods (Backus, Routledge, and Zin, 2005). The recursive preference allows the separation of coefficient of relative risk aversion γ and intertemporal elasticity of substitution (IES) ψ, which can be simultaneously large, whereas in standard CRRA preference $\psi = 1/\gamma$. This feature is desirable as an individual's willingness to take financial risk, does not have to be necessarily associated with the inverse of her inclination to substitute today's consumption with future consumption in response to change in intertemporal prices (Chen, Favilukis, and Ludvigson, 2011). When IES is larger than $1/\gamma$, agents prefer early resolution of uncertainty and exposure to the long-run risk carries high risk premium.

Hansen and Sargent (2010) consider the model in which agents are concerned with model mis-specification and Bayesian updates the parameters of several different sub-models. The agent will also update model-mixing probability and decide which sub-model is the most likely. This setup generates elevated risk premium since agent's distrust of her model adds to the price of risk.

Croce, Lettau, and Ludvigson (2012) investigates the role of information in consumption based long-run risk model and is able to explain the downward slope of equity term structure. When investors can identify both the short-run and long-run components of consumption risk, the standard long-run risk model can generate a sizable equity market risk premium only if the equity term structure slopes up. However, when investors cannot distinguish short-run and long-run components of consumption risk, the model is able to generate both a large equity market risk premium and a downward sloping equity term structure, which is what we observe in data.

Bansal and Shaliastovich (2011) focuses on the endogenous choice in learning. The agents may either pay a cost to learn the true long-run component or use Kalman filtering based on historical data. The actions of investors to learn about the true state can explain the asset-price jumps. The model implies income volatility can predict future jumps in returns.

Ai (2010) studies the implication of public information quality about persistent productivity shocks in a model with Kreps-Porteus preferences. The production-based long-run risk model with learning implies that when information quality is low equity premium is high and volatility of the risk-free rate is low.

Chapter 3

Model Setup

3.1 Preferences

There are two countries in the economy which are denoted home and foreign. Markets are complete. As in Colacito and Croce (2011), perfect home bias is imposed in the setup so that representative agent of each country would only consume the goods that she is endowed with. Since markets are complete, agents can still trade Arrow-Debreu securities and they have no home bias over financial assets. The first order conditions with respect to the purchase of those financial assets (No-Arbitrage equations) pin down the value of all assets and the growth rate of the exchange rate. For compactness of representation, only home's preferences and prices are characterized in this section. The foreign counterparts are denoted with identical expression with a superscript "*" added to designate the variables as foreign.

Following the long-run risk literature which emphasizes the importance of long-run economic growth prospects (Bansal and Yaron, 2004), the home representative agent has the Epstein-Zin-Weil (1989) preference:

$$U_t = \left\{(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(E_t\left[U_{t+1}^{1-\gamma}\right]\right)^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}$$

A noteworthy feature of the Epstein-Zin-Weil preference is the separation of the coefficient of intertemporal elasticity of substitution (IES) ψ and the coefficient of relative risk aversion γ , which can be simultaneously larger than one; whereas in standard CRRA preference $\psi = 1/\gamma$.

When $IES > 1/\gamma$, agents prefer early resolution of uncertainty(Epstein and Zin, 1991). For convenience, we can define composite parameter $\theta = \frac{1-\gamma}{1-1/\psi}$. Denote the ex-dividend price dividend ratio of an asset that pays a consumption stream C_t at end of period t as $W_{c,t} = P_t^C/C_t$. Also denote the ex-dividend price dividend ratio of an asset that pays a consumption stream D_t at end of period t as $W_{d,t} = P_t^D/D_t$. M_{t+1} is the pricing kernel and $R_{f,t}$ is the return on a one-period risk-free asset at time t. Given the Epstein-Zin-Weil preference, the optimal consumption choice yields the following asset pricing equations:

$$
\begin{aligned}
E_t[M_{t+1}R_{c,t+1}] = 1 \qquad & R_{c,t+1} = (P_{t+1}^C + C_{t+1})/P_t^C \\
E_t[M_{t+1}R_{d,t+1}] = 1 \qquad & R_{d,t+1} = (P_{t+1}^D + D_{t+1})/P_t^D
\end{aligned}
\tag{3.1}
$$

$$
\begin{aligned}
M_{t+1} &= \left(\delta\left[\frac{C_{t+1}}{C_t}\right]^{-1/\psi}\right)^{\theta} R_{c,t+1}^{\theta-1} \\
R_{f,t} &= (E_t[M_{t+1}])^{-1}
\end{aligned}
\tag{3.2}
$$

One can show the log pricing kernel is a function of log consumption growth Δc_t, and the log return on an asset which pays the consumption stream $r_{c,t+1}$:

$$
m_{t+1} = \frac{1-\gamma}{1-1/\psi}log\delta - \frac{1-\gamma}{\psi-1}\Delta c_{t+1} + \frac{1/\psi-\gamma}{1-1/\psi}r_{c,t+1} \tag{3.3}
$$

The Epstein-Zin-Weil preference and its basic asset pricing implications are shared by the three information structures describe below.

3.2 Full Information

Let the lower-case letters denote the variables in logarithms. The processes for log consumption and log dividend growth in the case of full information are specified as follows:

$$\Delta c_t = \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \tag{3.4}$$
$$\Delta c_t^* = \mu + z_t^* + \sigma_{c,f} \cdot \varepsilon_{c,t}^*$$
$$z_t = \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t}$$
$$z_t^* = \rho \cdot z_{t-1}^* + \sigma_{z,f} \cdot \varepsilon_{z,t}^*$$
$$\Delta d_t = \mu_d + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t}$$
$$\Delta d_t^* = \mu_d + \lambda \cdot z_t^* + \sigma_{d,f} \cdot \varepsilon_{d,t}^*$$

$$(\varepsilon_{c,t}, \varepsilon_{c,t}^*, \varepsilon_{z,t}, \varepsilon_{z,t}^*, \varepsilon_{d,t}, \varepsilon_{d,t}^*) \sim N.i.i.d(0, \Omega_c)$$

$$\Omega_c = \begin{bmatrix} 1 & \rho_c & 0 & 0 & 0 & 0 \\ \rho_c & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \rho_z & 0 & 0 \\ 0 & 0 & \rho_z & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \rho_d \\ 0 & 0 & 0 & 0 & \rho_d & 1 \end{bmatrix}$$

The small persistent predictable components z_t, z_t^* capture the small but persistent component of expected consumption and dividend growth rate in home and foreign countries, respectively. Note the timing convention here is that the predictive component z_t enters the consumption growth at time t rather than time t+1. In Bansal and Yaron (2004), z_t is the conditional expectation of consumption of next period, whereas in this timing convention the conditional expectation of consumption of next period is ρz_t. Let $\varepsilon_{c,t}, \varepsilon_{c,t}^*$ denote the short-run shocks to home and foreign consumption. $\varepsilon_{z,t}, \varepsilon_{z,t}^*$ are the long-run shocks to the highly persistent component. They are called the long-run shocks because due to the highly persistent nature of z_t, even a small innovation can impact cash flows for a very long time and thus agents demand a high risk premia for the long-run shocks when $IES > 1/\gamma$. $\varepsilon_{d,t}, \varepsilon_{d,t}^*$ are the dividend shocks and λ can be interpreted as the leverage ratio (Abel, 1999). As in Bansal and

Yaron (2004), dividend is not exposed to short-run shock. Indeed, for the short-run, long-run, dividend shocks within one country, $\varepsilon_{c,t}, \varepsilon_{z,t}, \varepsilon_{d,t}$, each type of shock is orthogonal to any other type of shock. On the other hand, the cross-country correlations short-run, long-run, dividend shocks are denoted as ρ_c, ρ_z, and ρ_d, respectively.

In the full information case, agents observe all variables of interest and fully take equation 3.4 into account when making asset pricing decisions.

3.3 Limited Information-Learning from Consumption

Under the information structure of limited information with learning from consumption stream, agents are assumed to observe all historical data of consumption growth, however the agents do not observe the underlying long-run persistent components z_t and z_t^*. As a consequence of limited information, the agents do not know if the innovation in consumption is due to $\varepsilon_{c,t}$ or $\varepsilon_{z,t}$. Because the linear state space of equation 3.4 is jointly Gaussian, the agents can use the Kalman filter (Harvey, 1989) to infer the underlying latent variables, namely z_t and z_t^*. The Kalman filter is the optimal linear filter for Gaussian systems. Due to the recursive nature of the Kalman filter, it can also be viewed as the simplest dynamic Bayesian Network (Haykin, 2001; Murphy, 2002).

Under this information structure the agent is only learning from consumption stream but not dividend stream, the case of learning from both consumption and dividend stream would be investigated in the next section. It would still be interesting to price dividend under the current scenario, however, a question arises as to how to specify the exogenous dividend growth process. Suppose the dividend process is as in equation 3.4, then the agent must be able to infer the latent variables not only from the consumption stream, but also the dividend stream which contains additional information regarding true values of the latent variables. In order to achieve internal consistency with the information structure, dividend processes equation 3.4 is replaced with the following:

$$\Delta d_t = \mu_d + \lambda \cdot (\Delta c_t - \mu) \tag{3.5}$$
$$\Delta d_t^* = \mu_d + \lambda \cdot (\Delta c_t^* - \mu)$$

The setup of the exogenous endowment processes are otherwise identical to the full information case. Note that the dividend process does not add any additional information for the filtering as it is a linear transformation of the consumption process, thus the agent only updates the filtered latent variable in response to consumption innovation.

As in Bansal and Shaliastovich (2011), I assume that $\Omega_c$is known by the agent. In addition to knowledge about the structure of cash flow processes, agents of both countries share the same information set at time t which is:

$$I_t^c = \left\{ \{\Delta c_{t-i}\}_{\, i=0,1,\ldots}, \left\{\Delta c_{t-i}^*\right\}_{\, i=0,1,\ldots}, \Omega_c \right\}$$

Note that Ω_c is not time-varying. In fact, learning with off-steady-state Kalman filter can generate endogenous time-varying volatility; more details on this and the Forward Premium Anomaly can be found in section 6.2.

Let the filtered state vector for the case of learning from consumption stream be denoted by $\widehat{z}_t$, $\widehat{z_t^*}$

$$\widehat{z}_t = E_t[z_t | I_t^c] \tag{3.6}$$
$$\widehat{z_t^*} = E_t[z_t^* | I_t^c]$$

The variance covariance matrix of the filtering errors is:

$$P_t^c = \begin{bmatrix} Phh_t^c & Phf_t^c \\ Phf_t^c & Pff_t^c \end{bmatrix}$$

$$
\begin{aligned}
Phh_t^c &= E_t[(z_t - \widehat{z}_t)^2 | I_t^c] \\
Phf_t^c &= E_t[(z_t - \widehat{z}_t)(z_t^* - \widehat{z_t^*}) | I_t^c] \\
Pff_t^c &= E_t[(z_t^* - \widehat{z_t^*})^2 | I_t^c]
\end{aligned}
\tag{3.7}
$$

Given the filtered latent variables, the structure of the cash flow process implies the following innovation representation:

$$
\begin{aligned}
\Delta c_t &= \mu + \widehat{z}_t + \nu_{c,t} \\
\Delta c_t^* &= \mu + \widehat{z_t^*} + \nu_{c,t}^* \\
\Delta d_t &= \mu_d + \lambda \cdot (\Delta c_t - \mu) \\
\Delta d_t^* &= \mu_d + \lambda \cdot (\Delta c_t^* - \mu)
\end{aligned}
\tag{3.8}
$$

where the innovations in consumption in home and foreign countries are defined as:

$$
\begin{cases}
\nu_{c,t} = \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \widehat{z}_t \\
\nu_{c,t}^* = \sigma_{c,f} \cdot \varepsilon_{c,t}^* + z_t^* - \widehat{z_t^*}
\end{cases}
\tag{3.9}
$$

The filtering problem and its solutions are solved in section 4.1.

3.4 Limited Information-Learning from Consumption and Dividend

The setup of the exogenous endowment processes are identical to the full information case. Notably the dividend process is :

$$\Delta d_t = \mu_d + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t} \tag{3.10}$$

$$\Delta d_t^* = \mu_d + \lambda \cdot z_t^* + \sigma_{d,f} \cdot \varepsilon_{d,t}^*$$

Both the consumption process and the dividend process provide information for the filtering of the state vector. The agents use Kalman filter to update the filtered latent variables in response to both consumption innovation and dividend innovation. Agents of both countries share the same information set at time t which is:

$$I_t^d = \left\{ \{\Delta c_{t-i}\}_{\, i=0,1,\ldots}, \left\{\Delta c_{t-i}^*\right\}_{\, i=0,1,\ldots}, \{\Delta d_{t-i}\}_{\, i=0,1,\ldots}, \left\{\Delta d_{t-i}^*\right\}_{\, i=0,1,\ldots}, \Omega_c \right\} \tag{3.11}$$

Let the filtered state vector for the case of learning from both consumption and dividend stream be denoted by $\widetilde{z}_t$, $\widetilde{z_t^*}$

$$\widetilde{z}_t, = E_t[z_t | I_t^d] \tag{3.12}$$

$$\widetilde{z_t^*} = E_t[z_t^* | I_t^d]$$

The variance covariance matrix of the filtering errors is:

$$P_t^d = \begin{bmatrix} Phh_t^d & Phf_t^d \\ Phf_t^d & Pff_t^{cd} \end{bmatrix}$$

$$Phh_t^d = E_t[(z_t - \widetilde{z}_t)^2 | I_t^d] \tag{3.13}$$

$$Phf_t^d = E_t[(z_t - \widetilde{z}_t)(z_t^* - \widetilde{z_t^*}) | I_t^d]$$

$$Pff_t^d = E_t[(z_t^* - \widetilde{z_t^*})^2 | I_t^d]$$

Given the filtered latent variables, the structure of the cash flow process implies the following innovation representation:

$$\Delta c_t = \mu + \widetilde{z}_t + \nu_{c,t} \tag{3.14}$$
$$\Delta c_t^* = \mu + \widetilde{z}_t^* + \nu_{c,t}^*$$
$$\Delta d_t = \mu_d + \lambda \cdot \widetilde{z}_t + \nu_{d,t}$$
$$\Delta d_t^* = \mu_d + \lambda \cdot \widetilde{z}_t^* + \nu_{d,t}^*$$

where the innovations in consumption and dividend, in home and foreign countries are defined as:

$$\begin{cases} \nu_{c,t} = \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \widetilde{z}_t \\ \nu_{c,t}^* = \sigma_{c,f} \cdot \varepsilon_{c,t}^* + z_t^* - \widetilde{z}_t^* \\ \nu_{d,t} = \sigma_{d,h} \cdot \varepsilon_{d,t} + \lambda \cdot (z_t - \widetilde{z}_t) \\ \nu_{d,t}^* = \sigma_{d,f} \cdot \varepsilon_{d,t}^* + \lambda \cdot \left(z_t^* - \widetilde{z}_t^*\right) \end{cases} \tag{3.15}$$

The filtering problem and its solutions are solved in section 4.2

Chapter 4

Derivations

4.1 Kalman Filter Derivation-Learning from Consumption

The consumption and dividend processes are assumed to be the following:

$$\Delta c_t = \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \tag{4.1}$$

$$\Delta c_t^* = \mu + z_t^* + \sigma_{c,f} \cdot \varepsilon_{c,t}^*$$

$$z_t = \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t} \tag{4.2}$$

$$z_t^* = \rho \cdot z_{t-1}^* + \sigma_{z,f} \cdot \varepsilon_{z,t}^* \tag{4.3}$$

$$\Delta d_t = \mu_d + \lambda \cdot (z_t + \sigma_{c,h} \cdot \varepsilon_{c,t})$$

$$\Delta d_t^* = \mu_d + \lambda \cdot \left(z_t^* + \sigma_{c,f} \cdot \varepsilon_{c,t}^*\right)$$

$$Var\begin{pmatrix} \sigma_{c,h} \cdot \varepsilon_{c,t} \\ \sigma_{c,f} \cdot \varepsilon_{c,t}^* \\ \sigma_{z,h} \cdot \varepsilon_{z,t} \\ \sigma_{z,f} \cdot \varepsilon_{z,t}^* \end{pmatrix} = \begin{bmatrix} \sigma_{c,h}^2 & \rho_c \cdot \sigma_{c,h}\sigma_{c,f} & 0 & 0 \\ \rho_c \cdot \sigma_{c,h}\sigma_{c,f} & \sigma_{c,f}^2 & 0 & 0 \\ 0 & 0 & \sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h}\sigma_{z,f} \\ 0 & 0 & \rho_z \cdot \sigma_{z,h}\sigma_{z,f} & \sigma_{z,f}^2 \end{bmatrix}$$

Define:

$$A^c = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \qquad H^c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q^c = \begin{bmatrix} \sigma_{c,h}^2 & \rho_c \cdot \sigma_{c,h}\sigma_{c,f} \\ \rho_c \cdot \sigma_{c,h}\sigma_{c,f} & \sigma_{c,f}^2 \end{bmatrix} \qquad R^c = \begin{bmatrix} \sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h}\sigma_{z,f} \\ \rho_z \cdot \sigma_{z,h}\sigma_{z,f} & \sigma_{z,f}^2 \end{bmatrix}$$

The state vector is (z_t, z_t^*), and the measurement vector is $(\Delta c_t, \Delta c_t^*)$. In addition, the filtered state vector for the case of learning from consumption process is $(\widehat{z_t}, \widehat{z_t^*})$.

The variance covariance matrix of the filtering errors is:

$$P_t^c = \begin{bmatrix} Phh_t^c & Phf_t^c \\ Phf_t^c & Pff_t^c \end{bmatrix}$$

Then by applying the standard Kalman filter update equation

$$P_t^c = A^c P_{t-1}^c A^{c'} - A^c P_{t-1}^c H^{c'} [H^c P_{t-1}^c H^{c'} + R^c]^{-1} H^c P_{t-1}^c A^{c'} + Q^c$$

One can show the one-step-ahead evolution equations for the variances of the filtering errors are:

$$Phh_t^c = \frac{\left(\rho^2 \sigma_{c,h}{}^2 \left(Phh_{t-1}^c \left(-\rho_c{}^2 \sigma_{c,f}{}^2 + Pff_{t-1}^c + \sigma_{c,f}{}^2\right) - Phf_{t-1}^c{}^2\right)\right)}{\left(\sigma_{c,f}{}^2 \left(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^c + \sigma_{c,h}{}^2\right) - 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} + Pff_{t-1}^c \left(Phh_{t-1}^c + \sigma_{c,h}{}^2\right) - Phf_{t-1}^c{}^2\right)} + \sigma_{z,h}{}^2 \tag{4.4}$$

$$Phf_t^c = \frac{\left(\rho^2 \sigma_{c,h}\sigma_{c,f} \left(\left(\rho_c{}^2 - 1\right) Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} + \rho_c \left(Phf_{t-1}^c{}^2 - Pff_{t-1}^c Phh_{t-1}^c\right)\right)\right)}{\left(-\sigma_{c,f}{}^2 \left(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^c + \sigma_{c,h}{}^2\right) + 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} - Pff_{t-1}^c \left(Phh_{t-1}^c + \sigma_{c,h}{}^2\right) + Phf_{t-1}^c{}^2\right)} + \rho_z \sigma_{z,f}\sigma_{z,h} \tag{4.5}$$

$$Pff_t^c = \frac{\left(\rho^2 \sigma_{c,f}{}^2 \left(\left(\rho_c{}^2 - 1\right) Pff_{t-1}^c \sigma_{c,h}{}^2 - Pff_{t-1}^c Phh_{t-1}^c + Phf_{t-1}^c{}^2\right)\right)}{\left(-\sigma_{c,f}{}^2 \left(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^c + \sigma_{c,h}{}^2\right) + 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} - Pff_{t-1}^c \left(Phh_{t-1}^c + \sigma_{c,h}{}^2\right) + Phf_{t-1}^c{}^2\right)} + \sigma_{z,f}{}^2 \tag{4.6}$$

One can derive the 2 by 2 Kalman gains in this case:

$$K^c_{,t} = \begin{bmatrix} K^c_{11,t} & K^c_{12,t} \\ K^c_{21,t} & K^c_{22,t} \end{bmatrix}$$

where

$$K^c_{11,t} = \frac{\left(Phf^c_{t-1}(-Phf^c_{t-1} - \rho_c \sigma_{c,h}\sigma_{c,f})\right) + \left(Phh^c_{t-1}\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right)\right)}{\left(\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right) - (Phf^c_{t-1} + \rho_c \sigma_{c,h}\sigma_{c,f})^2\right)} \tag{4.7}$$

$$K^c_{12,t} = \frac{\left(Phf^c_{t-1}\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\right) + \left(Phh^c_{t-1}(-Phf^c_{t-1} - \rho_c \sigma_{c,h}\sigma_{c,f})\right)}{\left(\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right) - (Phf^c_{t-1} + \rho_c \sigma_{c,h}\sigma_{c,f})^2\right)} \tag{4.8}$$

$$K^c_{21,t} = \frac{\left(Pff^c_{t-1}(-Phf^c_{t-1} - \rho_c \sigma_{c,h}\sigma_{c,f})\right) + \left(Phf^c_{t-1}\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right)\right)}{\left(\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right) - (Phf^c_{t-1} + \rho_c \sigma_{c,h}\sigma_{c,f})^2\right)} \tag{4.9}$$

$$K^c_{22,t} = \frac{\left(Pff^c_{t-1}\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\right) + \left(Phf^c_{t-1}(-Phf^c_{t-1} - \rho_c \sigma_{c,h}\sigma_{c,f})\right)}{\left(\left({\sigma_{c,h}}^2 + Phh^c_{t-1}\right)\left({\sigma_{c,f}}^2 + Pff^c_{t-1}\right) - (Phf^c_{t-1} + \rho_c \sigma_{c,h}\sigma_{c,f})^2\right)} \tag{4.10}$$

Or, in the innovation presentation, let $\nu_{LC,t} = \begin{pmatrix} \nu_{c,t} \\ \nu^*_{c,t} \end{pmatrix}$

The innovations in consumption in home and foreign countries are defined as:

$$\begin{cases} \nu_{c,t} = \sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \widehat{z}_t \\ \nu^*_{c,t} = \sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \widehat{z^*_t} \end{cases} \tag{4.11}$$

We have,

$$\Delta c_t = \mu + \widehat{z}_t + \nu_{c,t} \tag{4.12}$$

$$\Delta c^*_t = \mu + \widehat{z^*_t} + \nu^*_{c,t} \tag{4.13}$$

$$\Delta d_t = \mu_d + \lambda \cdot (\Delta c_t - \mu) \tag{4.14}$$

$$\Delta d^*_t = \mu_d + \lambda \cdot (\Delta c^*_t - \mu) \tag{4.15}$$

And the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:

$$\widehat{z_t} = \rho \cdot \widehat{z_{t-1}} + K^c_{11,t} \cdot \nu_{c,t} + K^c_{12,t} \cdot \nu^*_{c,t} \tag{4.16}$$

$$\widehat{z^*_t} = \rho \cdot \widehat{z^*_{t-1}} + K^c_{21,t} \cdot \nu_{c,t} + K^c_{22,t} \cdot \nu^*_{c,t} \tag{4.17}$$

The steady state Kalman filter is the solution to the following Discrete Algebraic Riccati Equation:

$$A^c.P^c_{ss}.A^{c^T} - A^c.P^c_{ss}.H^{c^T}.\left[H^c.P^c_{ss}.H^{c^T} + R^c\right]^{-1}.H^c.P^c_{ss}.A^{c^T} + Q^c \tag{4.18}$$

4.2 Kalman Filter Derivation-Learning from Consumption and Dividend

The consumption and dividend processes are assumed to be the following:

$$
\begin{aligned}
\Delta c_t &= \mu + z_t + \sigma_{c,h} \cdot \varepsilon_{c,t} \\
\Delta c_t^* &= \mu + z_t^* + \sigma_{c,f} \cdot \varepsilon_{c,t}^* \\
z_t &= \rho \cdot z_{t-1} + \sigma_{z,h} \cdot \varepsilon_{z,t} \\
z_t^* &= \rho \cdot z_{t-1}^* + \sigma_{z,f} \cdot \varepsilon_{z,t}^* \\
\Delta d_t &= \mu_d + \lambda \cdot z_t + \sigma_{d,h} \cdot \varepsilon_{d,t} \\
\Delta d_t^* &= \mu_d + \lambda \cdot z_t^* + \sigma_{d,f} \cdot \varepsilon_{d,t}^*
\end{aligned}
\tag{4.19}
$$

$$
\Omega = Var \begin{pmatrix} \sigma_{c,h} \cdot \varepsilon_{c,t} \\ \sigma_{c,f} \cdot \varepsilon_{c,t}^* \\ \sigma_{z,h} \cdot \varepsilon_{z,t} \\ \sigma_{z,f} \cdot \varepsilon_{z,t}^* \\ \sigma_{d,h} \cdot \varepsilon_{d,t} \\ \sigma_{d,f} \cdot \varepsilon_{d,t}^* \end{pmatrix} =
\begin{bmatrix}
\sigma_{c,h}^2 & \rho_c \sigma_{c,h} \sigma_{c,f} & 0 & 0 & 0 & 0 \\
\rho_c \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2 & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma_{z,h}^2 & \rho_z \sigma_{z,h} \sigma_{z,f} & 0 & 0 \\
0 & 0 & \rho_z \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma_{d,h}^2 & \rho_d \sigma_{d,h} \sigma_{d,f} \\
0 & 0 & 0 & 0 & \rho_d \sigma_{d,h} \sigma_{d,f} & \sigma_{d,f}^2
\end{bmatrix}
$$

Define:

$$
A^d = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \qquad H^d = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \lambda & 0 \\ 0 & \lambda \end{bmatrix}
$$

$$
Q^d = \begin{bmatrix} \sigma_{c,h}^2 & \rho_c \cdot \sigma_{c,h} \sigma_{c,f} \\ \rho_c \cdot \sigma_{c,h} \sigma_{c,f} & \sigma_{c,f}^2 \end{bmatrix} \quad
R^d = \begin{bmatrix}
\sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} & 0 & 0 \\
\rho_z \cdot \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2 & 0 & 0 \\
0 & 0 & \sigma_{z,h}^2 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} \\
0 & 0 & \rho_z \cdot \sigma_{z,h} \sigma_{z,f} & \sigma_{z,f}^2
\end{bmatrix}
$$

The state vector is $\begin{pmatrix} z_t \\ z_t^* \end{pmatrix}$.

The measurement vector is $\begin{pmatrix} \Delta c_t \\ \Delta c_t^* \\ \Delta d_t \\ \Delta d_t^* \end{pmatrix}$.

The filtered state vector for the case of learning from both consumption and dividend processes is $\begin{pmatrix} \widetilde{z}_t \\ z_t^* \end{pmatrix}$. Both the consumption process and the dividend process provide information for the filtering of the state vector.

The variance covariance matrix of the filtering errors is:

$$P_t^d = \begin{bmatrix} Phh_t^d & Phf_t^d \\ & \\ Phf_t^d & Pff_t^{cd} \end{bmatrix}$$

Then by applying the standard Kalman filter update equation

$$P_t^d = A^d P_{t-1}^d A^{d'} - A^d P_{t-1}^d H^{d'} [H^d P_{t-1}^d H^{d'} + R^d]^{-1} H^d P_{t-1}^d A^{d'} + Q^d$$

One can show the one-step-ahead evolution equations for the variances of the filtering errors are:

$$\begin{aligned}
Phh_t^d = \Big[& Pff_{t-1}^d \Big(Phh_{t-1}^d \rho^2 {\sigma_{d,h}}^2 {\sigma_{c,h}}^2 \left(-\left({\rho_c}^2 - 1\right) \lambda^2 {\sigma_{c,f}}^2 - \left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 \right) \\
& + {\sigma_{z,h}}^2 \lambda^2 \Big({\sigma_{d,h}}^2 {\sigma_{c,f}}^2 \Big(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^d + {\sigma_{c,h}}^2 \Big) - 2\rho_c \rho_d Phh_{t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} + Phh_{t-1}^d {\sigma_{d,f}}^2 {\sigma_{c,h}}^2 \Big) \\
& + {\sigma_{z,h}}^2 \Big(-\left({\rho_c}^2 - 1\right) \lambda^4 Phh_{t-1}^d {\sigma_{c,h}}^2 {\sigma_{c,f}}^2 - \left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 {\sigma_{d,h}}^2 \Big(Phh_{t-1}^d + {\sigma_{c,h}}^2 \Big) \Big) \Big) \\
& + {Phf_{t-1}^d}^2 \left(\rho^2 {\sigma_{d,h}}^2 {\sigma_{c,h}}^2 \left(\left({\rho_c}^2 - 1\right) \lambda^2 {\sigma_{c,f}}^2 + \left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 \right) \right. \\
& + {\sigma_{z,h}}^2 \left(\left({\rho_c}^2 - 1\right) \lambda^4 {\sigma_{c,h}}^2 {\sigma_{c,f}}^2 - \lambda^2 \left(-2\rho_c \rho_d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} + {\sigma_{d,f}}^2 {\sigma_{c,h}}^2 + {\sigma_{d,h}}^2 {\sigma_{c,f}}^2 \right) + \left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 {\sigma_{d,h}}^2 \right) \Big) \\
& + 2 Phf_{t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} {\sigma_{z,h}}^2 \left(\left({\rho_c}^2 - 1\right) \rho_d \lambda^2 \sigma_{c,h} \sigma_{c,f} + \rho_c \left({\rho_d}^2 - 1\right) \sigma_{d,f} \sigma_{d,h} \right) \\
& + {\sigma_{d,f}}^2 {\sigma_{c,f}}^2 \Big(\left({\rho_c}^2 - 1\right) \left({\rho_d}^2 - 1\right) Phh_{t-1}^d \rho^2 {\sigma_{d,h}}^2 {\sigma_{c,h}}^2 \\
& + {\sigma_{z,h}}^2 \Big(-\left({\rho_d}^2 - 1\right) {\sigma_{d,h}}^2 \Big(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^d + {\sigma_{c,h}}^2 \Big) - \left({\rho_c}^2 - 1\right) \lambda^2 Phh_{t-1}^d {\sigma_{c,h}}^2 \Big) \Big) \Big] \\
\div \Big[& -\left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 {\sigma_{d,h}}^2 \Big({\sigma_{c,f}}^2 \Big(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^d + {\sigma_{c,h}}^2 \Big) + Pff_{t-1}^d \Big(Phh_{t-1}^d + {\sigma_{c,h}}^2 \Big) \Big) \\
& + {Phf_{t-1}^d}^2 \left(\left({\rho_c}^2 - 1\right) \lambda^4 {\sigma_{c,h}}^2 {\sigma_{c,f}}^2 - \lambda^2 \left(-2\rho_c \rho_d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} + {\sigma_{d,f}}^2 {\sigma_{c,h}}^2 + {\sigma_{d,h}}^2 {\sigma_{c,f}}^2 \right) \right. \\
& \left. + \left({\rho_d}^2 - 1\right) {\sigma_{d,f}}^2 {\sigma_{d,h}}^2 \right) + 2 Phf_{t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} \left(\left({\rho_c}^2 - 1\right) \rho_d \lambda^2 \sigma_{c,h} \sigma_{c,f} + \rho_c \left({\rho_d}^2 - 1\right) \sigma_{d,f} \sigma_{d,h} \right) \\
& + \lambda^2 \Big({\sigma_{c,f}}^2 \Big(Pff_{t-1}^d Phh_{t-1}^d {\sigma_{d,h}}^2 - \left({\rho_c}^2 - 1\right) {\sigma_{c,h}}^2 \Big(Pff_{t-1}^d {\sigma_{d,h}}^2 + Phh_{t-1}^d {\sigma_{d,f}}^2 \Big) \Big) \\
& - 2\rho_c \rho_d Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f} \sigma_{d,h} \sigma_{c,h} \sigma_{c,f} + Pff_{t-1}^d Phh_{t-1}^d {\sigma_{d,f}}^2 {\sigma_{c,h}}^2 \Big) - \left({\rho_c}^2 - 1\right) \lambda^4 Pff_{t-1}^d Phh_{t-1}^d {\sigma_{c,h}}^2 {\sigma_{c,f}}^2 \Big]
\end{aligned}$$

$$
\begin{aligned}
Phf_t^d = \Big[& \rho_z \sigma_{z,f} \sigma_{z,h} \Big(-\left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \Big(\sigma_{c,f}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) + Pff_{t-1}^d \Big(Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) \Big) \\
& + Phf_{t-1}^d{}^2 \left(\left(\rho_c{}^2-1\right) \lambda^4 \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 - \lambda^2 \left(-2\rho_c\rho_d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + \sigma_{d,f}{}^2\sigma_{c,h}{}^2 + \sigma_{d,h}{}^2\sigma_{c,f}{}^2 \right) + \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \right) \\
& + 2Phf_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} \left(\left(\rho_c{}^2-1\right) \rho_d \lambda^2 \sigma_{c,h}\sigma_{c,f} + \rho_c \left(\rho_d{}^2-1\right) \sigma_{d,f}\sigma_{d,h} \right) \\
& + \lambda^2 \sigma_{c,f}{}^2 \Big(Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,h}{}^2 - \left(\rho_c{}^2-1\right) \sigma_{c,h}{}^2 \Big(Pff_{t-1}^d \sigma_{d,h}{}^2 + Phh_{t-1}^d \sigma_{d,f}{}^2 \Big) \Big) \\
& + \lambda^2 \Big(-2\rho_c\rho_d Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}{}^2 \sigma_{c,h}{}^2 \Big) \\
& - \left(\rho_c{}^2-1\right) \lambda^4 Pff_{t-1}^d Phh_{t-1}^d \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 \Big) \\
& + \rho^2 \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} \Big(\left(\rho_d{}^2-1\right) \sigma_{d,f}\sigma_{d,h} \Big(\left(\rho_c{}^2-1\right) Phf_{t-1}^d \sigma_{c,h}\sigma_{c,f} + \rho_c \Big(Phf_{t-1}^d{}^2 - Pff_{t-1}^d Phh_{t-1}^d \Big) \Big) \\
& + \left(\rho_c{}^2-1\right) \rho_d \lambda^2 \sigma_{c,h}\sigma_{c,f} \Big(Phf_{t-1}^d{}^2 - Pff_{t-1}^d Phh_{t-1}^d \Big) \Big) \Big] \\
\div \Big[& - \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \Big(\sigma_{c,f}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) + Pff_{t-1}^d \Big(Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) \Big) \\
& + Phf_{t-1}^d{}^2 \left(\left(\rho_c{}^2-1\right) \lambda^4 \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 - \lambda^2 \left(-2\rho_c\rho_d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + \sigma_{d,f}{}^2\sigma_{c,h}{}^2 + \sigma_{d,h}{}^2\sigma_{c,f}{}^2 \right) \right. \\
& \left. + \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \right) + 2Phf_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} \left(\left(\rho_c{}^2-1\right) \rho_d \lambda^2 \sigma_{c,h}\sigma_{c,f} + \rho_c \left(\rho_d{}^2-1\right) \sigma_{d,f}\sigma_{d,h} \right) \\
& + \lambda^2 \Big(\sigma_{c,f}{}^2 \Big(Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,h}{}^2 - \left(\rho_c{}^2-1\right) \sigma_{c,h}{}^2 \Big(Pff_{t-1}^d \sigma_{d,h}{}^2 + Phh_{t-1}^d \sigma_{d,f}{}^2 \Big) \Big) \\
& - 2\rho_c\rho_d Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}{}^2 \sigma_{c,h}{}^2 \Big) - \left(\rho_c{}^2-1\right) \lambda^4 Pff_{t-1}^d Phh_{t-1}^d \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 \Big]
\end{aligned}
$$

$$
\begin{aligned}
Pff_t^d = \Big[& Pff_{t-1}^d \Big(\rho^2 \sigma_{d,f}{}^2 \sigma_{c,f}{}^2 \Big(-\left(\rho_d{}^2-1\right) \sigma_{d,h}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) - \left(\rho_c{}^2-1\right) \lambda^2 Phh_{t-1}^d \sigma_{c,h}{}^2 \Big) \\
& + \sigma_{z,f}{}^2 \lambda^2 \Big(\sigma_{d,h}{}^2 \sigma_{c,f}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) - 2\rho_c\rho_d Phh_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + Phh_{t-1}^d \sigma_{d,f}{}^2 \sigma_{c,h}{}^2 \Big) \\
& + \sigma_{z,f}{}^2 \Big(-\left(\rho_c{}^2-1\right) \lambda^4 Phh_{t-1}^d \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 - \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \Big(Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) \Big) \Big) \\
& + Phf_{t-1}^d{}^2 \left(\rho^2 \sigma_{d,f}{}^2 \sigma_{c,f}{}^2 \left(\left(\rho_c{}^2-1\right) \lambda^2 \sigma_{c,h}{}^2 + \left(\rho_d{}^2-1\right) \sigma_{d,h}{}^2 \right) \right. \\
& \left. + \sigma_{z,f}{}^2 \left(\left(\rho_c{}^2-1\right) \lambda^4 \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 - \lambda^2 \left(-2\rho_c\rho_d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + \sigma_{d,f}{}^2\sigma_{c,h}{}^2 + \sigma_{d,h}{}^2\sigma_{c,f}{}^2 \right) + \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \right) \right) \\
& + 2Phf_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\sigma_{z,f}{}^2 \left(\left(\rho_c{}^2-1\right) \rho_d \lambda^2 \sigma_{c,h}\sigma_{c,f} + \rho_c \left(\rho_d{}^2-1\right) \sigma_{d,f}\sigma_{d,h} \right) \\
& + \sigma_{d,f}{}^2 \sigma_{c,f}{}^2 \sigma_{z,f}{}^2 \Big(-\left(\rho_d{}^2-1\right) \sigma_{d,h}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) - \left(\rho_c{}^2-1\right) \lambda^2 Phh_{t-1}^d \sigma_{c,h}{}^2 \Big) \Big] \\
\div \Big[& - \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \Big(\sigma_{c,f}{}^2 \Big(-\rho_c{}^2 \sigma_{c,h}{}^2 + Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) + Pff_{t-1}^d \Big(Phh_{t-1}^d + \sigma_{c,h}{}^2 \Big) \Big) \\
& + Phf_{t-1}^d{}^2 \left(\left(\rho_c{}^2-1\right) \lambda^4 \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 - \lambda^2 \left(-2\rho_c\rho_d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + \sigma_{d,f}{}^2\sigma_{c,h}{}^2 + \sigma_{d,h}{}^2\sigma_{c,f}{}^2 \right) \right. \\
& \left. + \left(\rho_d{}^2-1\right) \sigma_{d,f}{}^2 \sigma_{d,h}{}^2 \right) + 2Phf_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} \left(\left(\rho_c{}^2-1\right) \rho_d \lambda^2 \sigma_{c,h}\sigma_{c,f} + \rho_c \left(\rho_d{}^2-1\right) \sigma_{d,f}\sigma_{d,h} \right) \\
& + \lambda^2 \Big(\sigma_{c,f}{}^2 \Big(Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,h}{}^2 - \left(\rho_c{}^2-1\right) \sigma_{c,h}{}^2 \Big(Pff_{t-1}^d \sigma_{d,h}{}^2 + Phh_{t-1}^d \sigma_{d,f}{}^2 \Big) \Big) \\
& - 2\rho_c\rho_d Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f} + Pff_{t-1}^d Phh_{t-1}^d \sigma_{d,f}{}^2 \sigma_{c,h}{}^2 \Big) - \left(\rho_c{}^2-1\right) \lambda^4 Pff_{t-1}^d Phh_{t-1}^d \sigma_{c,h}{}^2 \sigma_{c,f}{}^2 \Big]
\end{aligned}
$$

One can derive the 2 by 4 Kalman gains in this case:

$$
K_{,t}^d = \begin{bmatrix} K_{11,t}^d & K_{12,t}^d & K_{13,t}^d & K_{14,t}^d \\ K_{21,t}^d & K_{22,t}^d & K_{23,t}^d & K_{24,t}^d \end{bmatrix}
$$

where

$$
\begin{aligned}
K^d_{11,t} = \Big[&\sigma_{d,h}\left(\left({\rho_d}^2-1\right)\sigma_{d,h}\left({Phf^d_{t-1}}^2+\rho_c\sigma_{c,h}\sigma_{c,f}Phf^d_{t-1}-Phh^d_{t-1}\left({\sigma_{c,f}}^2+Pff^d_{t-1}\right)\right){\sigma_{d,f}}^2\right.\\
&\left.+\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right)\sigma_{c,f}(\rho_c\rho_d\sigma_{d,f}\sigma_{c,h}-\sigma_{d,h}\sigma_{c,f})\lambda^2\right)\Big]\\
\div\Big[&-\left({\rho_c}^2-1\right)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\right.\\
&\left.+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\left({\rho_c}^2-1\right)\left(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\right){\sigma_{c,h}}^2\right){\sigma_{c,f}}^2\right)\lambda^2\\
&-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\left(\left(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\right){\sigma_{c,f}}^2+Pff^d_{t-1}\left({\sigma_{c,h}}^2+Phh^d_{t-1}\right)\right)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\left(\left({\rho_c}^2-1\right)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\left({\rho_d}^2-1\right)\sigma_{d,f}\sigma_{d,h}\right)\\
&+{Phf^d_{t-1}}^2\left(\left({\rho_c}^2-1\right){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\left({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\right)\lambda^2\right.\\
&\left.+\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{12,t} = \Big[&\sigma_{d,h}\sigma_{c,h}\left(\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right)(\rho_c\sigma_{d,h}\sigma_{c,f}-\rho_d\sigma_{d,f}\sigma_{c,h})\lambda^2\right.\\
&\left.-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2\sigma_{d,h}(Phf^d_{t-1}\sigma_{c,h}-\rho_c Phh^d_{t-1}\sigma_{c,f})\right)\Big]\div\Big[-\left({\rho_c}^2-1\right)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\right.\\
&\left.+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\left({\rho_c}^2-1\right)\left(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\right){\sigma_{c,h}}^2\right){\sigma_{c,f}}^2\right)\lambda^2\\
&-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\left(\left(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\right){\sigma_{c,f}}^2+Pff^d_{t-1}\left({\sigma_{c,h}}^2+Phh^d_{t-1}\right)\right)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\left(\left({\rho_c}^2-1\right)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\left({\rho_d}^2-1\right)\sigma_{d,f}\sigma_{d,h}\right)\\
&+{Phf^d_{t-1}}^2\left(\left({\rho_c}^2-1\right){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\left({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\right)\lambda^2\right.\\
&\left.+\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{13,t} = \Big[&\sigma_{c,h}\lambda\left(\left({\rho_c}^2-1\right)\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right)\sigma_{c,h}{\sigma_{c,f}}^2\lambda^2\right.\\
&-\sigma_{d,f}\left({\rho_c}^2-1\right)(Phh^d_{t-1}\sigma_{d,f}-\rho_d Phf^d_{t-1}\sigma_{d,h})\sigma_{c,h}{\sigma_{c,f}}^2\\
&\left.+\left(\sigma_{d,f}\rho_c\rho_d\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right)\sigma_{d,h}\sigma_{c,f}+\left(Pff^d_{t-1}Phh^d_{t-1}-{Phf^d_{t-1}}^2\right)\sigma_{d,f}\sigma_{c,h}\right)\right)\Big]\\
\div\Big[&-\left({\rho_c}^2-1\right)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\right.\\
&\left.+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\left({\rho_c}^2-1\right)\left(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\right){\sigma_{c,h}}^2\right){\sigma_{c,f}}^2\right)\lambda^2\\
&-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\left(\left(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\right){\sigma_{c,f}}^2+Pff^d_{t-1}\left({\sigma_{c,h}}^2+Phh^d_{t-1}\right)\right)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\left(\left({\rho_c}^2-1\right)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\left({\rho_d}^2-1\right)\sigma_{d,f}\sigma_{d,h}\right)\\
&+{Phf^d_{t-1}}^2\left(\left({\rho_c}^2-1\right){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\left({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\right)\lambda^2\right.\\
&\left.+\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{14,t} = \Big[&\sigma_{d,h}\sigma_{c,h}\Big(-\big({\rho_c}^2-1\big)(\rho_d Phh^d_{t-1}\sigma_{d,f}-Phf^d_{t-1}\sigma_{d,h})\sigma_{c,h}{\sigma_{c,f}}^2+\rho_c\Big({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\Big)\sigma_{d,h}\sigma_{c,f}\\
&+\rho_d\Big(Pff^d_{t-1}Phh^d_{t-1}-{Phf^d_{t-1}}^2\Big)\sigma_{d,f}\sigma_{c,h}\Big)\lambda\Big]\div\Big[\big({\rho_c}^2-1\big)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&-\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\\
&+\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\big({\rho_c}^2-1\big)\Big(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\Big){\sigma_{c,h}}^2\Big){\sigma_{c,f}}^2\Big)\lambda^2\\
&+\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\Big(\Big(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\Big){\sigma_{c,f}}^2+Pff^d_{t-1}\Big({\sigma_{c,h}}^2+Phh^d_{t-1}\Big)\Big)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\big(-\big({\rho_c}^2-1\big)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2-\rho_c\big({\rho_d}^2-1\big)\sigma_{d,f}\sigma_{d,h}\big)\\
&+{Phf^d_{t-1}}^2\big(-\big({\rho_c}^2-1\big){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4+\big({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\big)\lambda^2\\
&-\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{21,t} = \Big[&\sigma_{d,f}\sigma_{c,f}\Big(\big({\rho_d}^2-1\big)\sigma_{d,f}(\rho_c Pff^d_{t-1}\sigma_{c,h}-Phf^d_{t-1}\sigma_{c,f}){\sigma_{d,h}}^2\\
&+\Big({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\Big)(\rho_c\sigma_{d,f}\sigma_{c,h}-\rho_d\sigma_{d,h}\sigma_{c,f})\lambda^2\Big)\Big]\div\Big[-\big({\rho_c}^2-1\big)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&+\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\\
&+\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\big({\rho_c}^2-1\big)\Big(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\Big){\sigma_{c,h}}^2\Big){\sigma_{c,f}}^2\Big)\lambda^2\\
&-\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\Big(\Big(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\Big){\sigma_{c,f}}^2+Pff^d_{t-1}\Big({\sigma_{c,h}}^2+Phh^d_{t-1}\Big)\Big)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\big(\big({\rho_c}^2-1\big)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\big({\rho_d}^2-1\big)\sigma_{d,f}\sigma_{d,h}\big)\\
&+{Phf^d_{t-1}}^2\big(\big({\rho_c}^2-1\big){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\big({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\big)\lambda^2\\
&+\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{22,t} = \Big[&\sigma_{d,f}\Big(\big({\rho_d}^2-1\big)\sigma_{d,f}{\sigma_{d,h}}^2\Big({Phf^d_{t-1}}^2+\rho_c\sigma_{c,h}\sigma_{c,f}Phf^d_{t-1}-Pff^d_{t-1}\Big({\sigma_{c,h}}^2+Phh^d_{t-1}\Big)\Big)\\
&-\Big({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\Big)\sigma_{c,h}(\sigma_{d,f}\sigma_{c,h}-\rho_c\rho_d\sigma_{d,h}\sigma_{c,f})\lambda^2\Big)\Big]\\
\div\Big[&-\big({\rho_c}^2-1\big)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4\\
&+\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\\
&+\Big(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\big({\rho_c}^2-1\big)\Big(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\Big){\sigma_{c,h}}^2\Big){\sigma_{c,f}}^2\Big)\lambda^2\\
&-\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\Big(\Big(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\Big){\sigma_{c,f}}^2+Pff^d_{t-1}\Big({\sigma_{c,h}}^2+Phh^d_{t-1}\Big)\Big)\\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\big(\big({\rho_c}^2-1\big)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\big({\rho_d}^2-1\big)\sigma_{d,f}\sigma_{d,h}\big)\\
&+{Phf^d_{t-1}}^2\big(\big({\rho_c}^2-1\big){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\big({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\big)\lambda^2\\
&+\big({\rho_d}^2-1\big){\sigma_{d,f}}^2{\sigma_{d,h}}^2\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{23,t} = \Big[&\sigma_{d,f}\sigma_{c,f}\left(\rho_c\left({Phf^d_{t-1}}^2 - Pff^d_{t-1}Phh^d_{t-1}\right)\sigma_{d,f}\sigma_{c,h}\right. \\
&\left.+\left(\left({\rho_c}^2-1\right)(Phf^d_{t-1}\sigma_{d,f}-\rho_d Pff^d_{t-1}\sigma_{d,h}){\sigma_{c,h}}^2+\rho_d\left(Pff^d_{t-1}Phh^d_{t-1}-{Phf^d_{t-1}}^2\right)\sigma_{d,h}\right)\sigma_{c,f}\right)\lambda\Big] \\
\div\Big[&\left({\rho_c}^2-1\right)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4 \\
&-\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\right. \\
&\left.+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\left({\rho_c}^2-1\right)\left(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\right){\sigma_{c,h}}^2\right){\sigma_{c,f}}^2\right)\lambda^2 \\
&+\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\left(\left(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\right){\sigma_{c,f}}^2+Pff^d_{t-1}\left({\sigma_{c,h}}^2+Phh^d_{t-1}\right)\right) \\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\left(-\left({\rho_c}^2-1\right)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2-\rho_c\left({\rho_d}^2-1\right)\sigma_{d,f}\sigma_{d,h}\right) \\
&+{Phf^d_{t-1}}^2\left(-\left({\rho_c}^2-1\right){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4+\left({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\right)\lambda^2\right. \\
&\left.-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\right)\Big]
\end{aligned}
$$

$$
\begin{aligned}
K^d_{24,t} = \Big[&\sigma_{c,f}\lambda\left(\left({\rho_c}^2-1\right)\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right){\sigma_{c,h}}^2\sigma_{c,f}\lambda^2\right. \\
&+\rho_c\rho_d\left({Phf^d_{t-1}}^2-Pff^d_{t-1}Phh^d_{t-1}\right)\sigma_{d,f}\sigma_{d,h}\sigma_{c,h} \\
&\left.+\sigma_{d,h}\left(-\sigma_{d,h}{Phf^d_{t-1}}^2+\left({\rho_c}^2-1\right)(\rho_d Phf^d_{t-1}\sigma_{d,f}-Pff^d_{t-1}\sigma_{d,h}){\sigma_{c,h}}^2+Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,h}\right)\sigma_{c,f}\right)\Big] \\
\div\Big[&-\left({\rho_c}^2-1\right)Pff^d_{t-1}Phh^d_{t-1}{\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4 \\
&+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d Pff^d_{t-1}Phh^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}\right. \\
&\left.+\left(Pff^d_{t-1}Phh^d_{t-1}{\sigma_{d,h}}^2-\left({\rho_c}^2-1\right)\left(Phh^d_{t-1}{\sigma_{d,f}}^2+Pff^d_{t-1}{\sigma_{d,h}}^2\right){\sigma_{c,h}}^2\right){\sigma_{c,f}}^2\right)\lambda^2 \\
&-\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\left(\left(-{\rho_c}^2{\sigma_{c,h}}^2+{\sigma_{c,h}}^2+Phh^d_{t-1}\right){\sigma_{c,f}}^2+Pff^d_{t-1}\left({\sigma_{c,h}}^2+Phh^d_{t-1}\right)\right) \\
&+2Phf^d_{t-1}\sigma_{d,f}\sigma_{d,h}\sigma_{c,h}\sigma_{c,f}\left(\left({\rho_c}^2-1\right)\rho_d\sigma_{c,h}\sigma_{c,f}\lambda^2+\rho_c\left({\rho_d}^2-1\right)\sigma_{d,f}\sigma_{d,h}\right) \\
&+{Phf^d_{t-1}}^2\left(\left({\rho_c}^2-1\right){\sigma_{c,h}}^2{\sigma_{c,f}}^2\lambda^4-\left({\sigma_{d,f}}^2{\sigma_{c,h}}^2-2\rho_c\rho_d\sigma_{d,f}\sigma_{d,h}\sigma_{c,f}\sigma_{c,h}+{\sigma_{d,h}}^2{\sigma_{c,f}}^2\right)\lambda^2\right. \\
&\left.+\left({\rho_d}^2-1\right){\sigma_{d,f}}^2{\sigma_{d,h}}^2\right)\Big]
\end{aligned}
$$

Or, in the innovation presentation, let $\nu_{LCD,t} = \begin{pmatrix} \nu_{c,t} \\ \nu^*_{c,t} \\ \nu_{d,t} \\ \nu^*_{d,t} \end{pmatrix}$

The innovations in consumption and dividend, in home and foreign countries are defined as:

$$
\begin{cases}
\nu_{c,t} = \sigma_{c,h}\cdot\varepsilon_{c,t} + z_t - \widetilde{z}_t \\
\nu^*_{c,t} = \sigma_{c,f}\cdot\varepsilon^*_{c,t} + z^*_t - \widetilde{z}^*_t \\
\nu_{d,t} = \sigma_{d,h}\cdot\varepsilon_{d,t} + \lambda\cdot(z_t - \widetilde{z}_t) \\
\nu^*_{d,t} = \sigma_{d,f}\cdot\varepsilon^*_{d,t} + \lambda\cdot\left(z^*_t - \widetilde{z}^*_t\right)
\end{cases}
\tag{4.20}
$$

We have,

$$\Delta c_t = \mu + \widehat{z}_t + \nu_{c,t} \tag{4.21}$$

$$\Delta c^*_t = \mu + \widehat{z^*_t} + \nu^*_{c,t} \tag{4.22}$$

$$\Delta d_t = \mu_d + \lambda \cdot \widehat{z}_t + \nu_{d,t} \tag{4.23}$$

$$\Delta d^*_t = \mu_d + \lambda \cdot \widehat{z^*_t} + \nu^*_{d,t} \tag{4.24}$$

And the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:

$$\widetilde{z}_t = \rho \cdot \widetilde{z_{t-1}} + K^d_{11,t} \cdot \nu_{c,t} + K^d_{12,t} \cdot \nu^*_{c,t} + K^d_{13,t} \cdot \nu_{d,t} + K^d_{14,t} \cdot \nu^*_{d,t} \tag{4.25}$$

$$\widetilde{z^*_t} = \rho \cdot \widetilde{z^*_{t-1}} + K^d_{21,t} \cdot \nu_{c,t} + K^d_{22,t} \cdot \nu^*_{c,t} + K^d_{23,t} \cdot \nu_{d,t} + K^d_{24,t} \cdot \nu^*_{d,t} \tag{4.26}$$

The steady state Kalman filter is the solution to the following Discrete Algebraic Riccati Equation:

$$A^d.P^d_{ss}.A^{d^T} - A^d.P^d_{ss}.H^{d^T}.\left[H^d.P^d_{ss}.H^{d^T} + R^d\right]^{-1}.H^d.P^d_{ss}.A^{d^T} + Q^d \tag{4.27}$$

4.3 Derivation of Pricing Kernel

In this section I elaborate the asset pricing results obtained from log-linearization of the Epstein-Zin Utility. This exercise was performed to provide some intuition through analytical solutions. I used numerical third order approximations in the simulations and plots and did not use the analytical first order approximation results presented in this section.

4.3.1 Cash Flow Model

The baseline linear state space representation is

$$\begin{aligned}
\Delta c_t &= \mu + z_t + \eta_{c,t} \\
\Delta c_t^* &= \mu + z_t^* + \eta_{c,t}^* \\
z_t &= \rho \cdot z_{t-1} + \eta_{z,t} \\
z_t^* &= \rho \cdot z_{t-1}^* + \eta_{z,t}^* \\
\Delta d_t &= \mu_d + \lambda \cdot z_t + \eta_{d,t} \\
\Delta d_t^* &= \mu_d + \lambda \cdot z_t^* + \eta_{d,t}^*
\end{aligned} \tag{4.28}$$

where

$$\eta_t = \begin{pmatrix} \eta_{c,t} \\ \eta_{c,t}^* \\ \eta_{z,t} \\ \eta_{z,t}^* \\ \eta_{d,t} \\ \eta_{d,t}^* \end{pmatrix} \backsim N.i.i.d.(0, S) \tag{4.29}$$

4.3.2 Log-Linearization

I follow (Croce, Lettau, and Ludvigson, 2012). This implementation involves two-countries and the results are similar to the standard long-run risk literature, up to a difference in the timing convention in the long-run persistent component. Define the price dividend ratio of an asset that pays a consumption stream C_t at end of period t as $W_{c,t} = P_t^C / C_t$, $R_{c,t+1} = (P_{t+1}^C + C_{t+1})/P_t^C$, then the Campbell-Shiller log-linearization yields:

$$w_{c,t} = \overline{w_c} + \sum_{i=0}^{\infty} \kappa_c E_t[\Delta c_{t+1+i}] - \sum_{i=0}^{\infty} \kappa_c^i E_t[r_{c,t+1+i}] \tag{4.30}$$

$$\kappa_c = \frac{\overline{exp(w_c)}}{1 + \overline{exp(w_c)}} \tag{4.31}$$

The first order condition of the Epstein-Zin Utility yields:

$$m_{t+1} = \overline{m} - \frac{1}{\psi} z_{t+1} - \kappa_c \frac{\gamma - 1/\psi}{1 - \rho\kappa_c} \eta_{z,t+1} - \gamma \eta_{c,t+1} \tag{4.32}$$

$$r_{c,t+1} = \overline{r_c} + \frac{1}{\psi} z_{t+1} + \kappa_c \frac{1 - 1/\psi}{1 - \rho\kappa_c} \eta_{z,t+1} + \eta_{c,t+1} \tag{4.33}$$

$$r_{f,t} = \overline{r_f} + \frac{1}{\psi} z_{t+1} \tag{4.34}$$

$$w_{c,t} = \overline{w_c} + \frac{1 - 1/\psi}{1 - \kappa_c \rho} z_{t+1} \tag{4.35}$$

Define the price dividend ratio of an asset that pays a consumption stream D_t at end of period t as $W_{d,t} = P_t^D / D_t$, $R_{d,t+1} = (P_{t+1}^D + D_{t+1})/P_t^D$, then the Campbell-Shiller log linearization yields:

$$w_{d,t} = \overline{w_d} + \frac{\lambda - 1/\psi}{1 - \kappa_d \rho} z_{t+1} \tag{4.36}$$

$$\kappa_d = \frac{\overline{exp(w_d)}}{1 + \overline{exp(w_d)}} \tag{4.37}$$

$$r_{d,t+1} = \overline{r_d} + \frac{1}{\psi} z_{t+1} + \kappa_d \frac{\lambda - 1/\psi}{1 - \rho\kappa_d} \eta_{z,t+1} + \eta_{d,t+1} \tag{4.38}$$

Thus, in vector form, we have that for the home country:

$$m_{t+1} = \overline{m} - \frac{1}{\psi} z_{t+1} + \Gamma_m \eta_{t+1} \tag{4.39}$$

$$\Gamma_m = \begin{bmatrix} -\gamma & 0 & -\kappa_c \frac{\gamma - 1/\psi}{1-\kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \tag{4.40}$$

$$r_{c,t+1} = \overline{r_c} + \frac{1}{\psi} z_{t+1} + \Gamma_c \eta_{t+1} \tag{4.41}$$

$$\Gamma_c = \begin{bmatrix} 1 & 0 & \kappa_c \frac{1 - 1/\psi}{1-\kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \tag{4.42}$$

$$r_{d,t+1} = \overline{r_d} + \frac{1}{\psi} z_{t+1} + \Gamma_d \eta_{t+1} \tag{4.43}$$

$$\Gamma_d = \begin{bmatrix} 0 & 0 & \kappa_d \frac{\lambda - 1/\psi}{1-\kappa_d \rho} & 0 & 1 & 0 \end{bmatrix} \tag{4.44}$$

Similarly, for the foreign country:

$$m^*_{t+1} = \overline{m}^* - \frac{1}{\psi} z^*_{t+1} + \Gamma^*_m \eta_{t+1} \tag{4.45}$$

$$\Gamma^*_m = \begin{bmatrix} 0 & -\gamma & 0 & -\kappa_c \frac{\gamma - 1/\psi}{1-\kappa_c \rho} & 0 & 0 \end{bmatrix} \tag{4.46}$$

$$r^*_{c,t+1} = \overline{r_c}^* + \frac{1}{\psi} z^*_{t+1} + \Gamma^*_c \eta_{t+1} \tag{4.47}$$

$$\Gamma^*_c = \begin{bmatrix} 0 & 1 & 0 & \kappa_c \frac{1 - 1/\psi}{1-\kappa_c \rho} & 0 & 0 \end{bmatrix} \tag{4.48}$$

$$r^*_{d,t+1} = \overline{r_d}^* + \frac{1}{\psi} z^*_{t+1} + \Gamma^*_d \eta_{t+1} \tag{4.49}$$

$$\Gamma^*_d = \begin{bmatrix} 0 & 0 & 0 & \kappa_d \frac{\lambda - 1/\psi}{1-\kappa_d \rho} & 0 & 1 \end{bmatrix} \tag{4.50}$$

Since

$$E_t[r^{ex}_{c,t+1}] = -cov\left(m_{t+1} - E_t[m_{t+1}], r_{c,t+1} - E_t[r_{c,t+1}]\right) - \frac{1}{2} Var\left(r_{c,t+1} - E_t[r_{c,t+1}]\right) \tag{4.51}$$

$$E_t[r^{ex}_{d,t+1}] = -cov\left(m_{t+1} - E_t[m_{t+1}], r_{d,t+1} - E_t[r_{d,t+1}]\right) - \frac{1}{2} Var\left(r_{d,t+1} - E_t[r_{d,t+1}]\right) \tag{4.52}$$

We have

$$E_t[r^{ex}_{c,t+1}] = -\Gamma_m S \Gamma^{'}_c - \frac{1}{2} \Gamma_c S \Gamma^{'}_c \tag{4.53}$$

$$E_t[r^{ex*}_{c,t+1}] = -\Gamma^*_m S \Gamma^{*'}_c - \frac{1}{2} \Gamma^*_c S \Gamma^{*'}_c \tag{4.54}$$

$$E_t[r^{ex}_{d,t+1}] = -\Gamma_m S \Gamma^{'}_d - \frac{1}{2} \Gamma_d S \Gamma^{'}_d \tag{4.55}$$

$$E_t[r^{ex*}_{d,t+1}] = -\Gamma^*_m S \Gamma^{*'}_d - \frac{1}{2} \Gamma^*_d S \Gamma^{*'}_d \tag{4.56}$$

4.3.3 Coefficients

By definition of the pricing kernel:

$$E[r_f] = -log\delta + \frac{1}{\psi}\mu + \frac{1-\theta}{\theta}\left(-\Gamma_m S\Gamma_c^{'} - \frac{1}{2}\Gamma_c S\Gamma_c^{'}\right) - \frac{1}{2\theta}\Gamma_m S\Gamma_m^{'} \tag{4.57}$$

where $\theta = \frac{1-\gamma}{1-1/\psi}$

Thus, the intercept can be shown to be

$$\overline{m} = \theta log\delta - \frac{\theta}{\psi}\mu + (\theta - 1)(E[r_c^{ex}] + E[r_f]) \tag{4.58}$$

Since the Euler Equations holds for all values of the long-run persistent component, plug-in the case $z_{t+1} = 0$ can pin down expressions for κ_c and κ_d

$$\kappa_c = \delta e^{\left(1-\frac{1}{\psi}\right)\left(\mu - \frac{1}{2}(\gamma-1)Var\left[\eta_{c,t+1} + \frac{\kappa_c \eta_{z,t+1}}{1-\kappa_c\rho}\right]\right)} \tag{4.59}$$

$$\kappa_d = e^{\overline{m}+\mu+\frac{1}{2}Var[(\Gamma_m+\Gamma_d)\eta_{t+1}]} \tag{4.60}$$

4.4 Mapping of Information Structure

The asset pricing results obtained from log-linearization of Epstein-Zin preference shown in section 4.3 can be readily applied to different information structures, as long as they are expressed in terms of the baseline linear state space representation of equation (4.28). To achieve this, I derive the mappings from innovation space representation into the baseline linear state space representation of consumption, long-run component, and dividend shocks.

4.4.1 Full Information

It is trivial to transform the full information case into the baseline linear state space representation

$$\begin{pmatrix} \eta_{c,t} \\ \eta^*_{c,t} \\ \eta_{z,t} \\ \eta^*_{z,t} \\ \eta_{d,t} \\ \eta^*_{d,t} \end{pmatrix} = \begin{bmatrix} \sigma_{c,h} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{c,f} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{z,h} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{z,f} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{d,h} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{d,f} \end{bmatrix} \begin{pmatrix} \varepsilon_{c,t} \\ \varepsilon^*_{c,t} \\ \varepsilon_{z,t} \\ \varepsilon^*_{z,t} \\ \varepsilon_{d,t} \\ \varepsilon^*_{d,t} \end{pmatrix} \tag{4.61}$$

In other words

$$\eta_t = \Sigma_{FI}\varepsilon_t \tag{4.62}$$

$$S_{FI} = \Omega \tag{4.63}$$

4.4.2 Learning from Consumption

The following mapping allows the transformation of the innovation space representation for the case of learning from consumption stream into the baseline linear state space representation.

$$\begin{pmatrix} \eta_{c,t} \\ \eta^*_{c,t} \\ \eta_{z,t} \\ \eta^*_{z,t} \\ \eta_{d,t} \\ \eta^*_{d,t} \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ K^c_{11,t} & K^c_{12,t} \\ K^c_{21,t} & K^c_{22,t} \\ \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} \nu_{c,t} \\ \nu^*_{c,t} \end{pmatrix} \tag{4.64}$$

In other words

$$\eta_t = \Sigma_{LC}\nu_{LC,t} \tag{4.65}$$

Thus

$$S_{LC} = \Sigma_{LC} P_{LC} \Sigma^{'}_{LC} \tag{4.66}$$

where K^c_{ij} are the elements Kalman Gain matrix for the case of learning from consumption stream. $P_{LC} = E[\nu_{LC}\nu^{'}_{LC}]$ is a non-linear transformation of the Ω matrix obtained by solving the steady state Kalman filtering problem.

4.4.3 Learning from Consumption and Dividend

The following mapping allows the transformation of the innovation space representation for the case of learning from consumption stream and dividend stream into the baseline linear state space representation.

$$\begin{pmatrix} \eta_{c,t} \\ \eta^*_{c,t} \\ \eta_{z,t} \\ \eta^*_{z,t} \\ \eta_{d,t} \\ \eta^*_{d,t} \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ K^d_{11,t} & K^d_{12,t} & K^d_{13,t} & K^d_{14,t} \\ K^d_{21,t} & K^d_{22,t} & K^d_{23,t} & K^d_{24,t} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \nu_{c,t} \\ \nu^*_{c,t} \\ \nu_{d,t} \\ \nu^*_{d,t} \end{pmatrix} \tag{4.67}$$

In other words

$$\eta_t = \Sigma_{LCD}\nu_{LCD,t} \tag{4.68}$$

Thus

$$S_{LCD} = \Sigma_{LCD} P_{LCD} \Sigma_{LCD}^{'} \tag{4.69}$$

where K_{ij}^{c} are the elements Kalman Gain matrix for the case of learning from consumption and dividend stream. $P_{LCD} = E[\nu_{LCD}\nu_{LCD}^{'}]$ is a non-linear transformation of the Ω matrix obtained by solving the steady state Kalman filtering problem.

Chapter 5

Analytical Model Solution

5.1 Full Information

The Euler equations 3.2 yield the policy function for price-consumption and price-dividend ratio as a function of the true long-run persistent components z_t, z_t^*. Section 4.3 derives the log-linearized solutions for full information. In vector form, we have that for the home country:

$$m_{t+1} = \overline{m} - \frac{1}{\psi} z_{t+1} + \Gamma_m \eta_{t+1} \tag{5.1}$$

$$\Gamma_m = \begin{bmatrix} -\gamma & 0 & -\kappa_c \frac{\gamma - 1/\psi}{1-\kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \tag{5.2}$$

$$r_{c,t+1} = \overline{r_c} + \frac{1}{\psi} z_{t+1} + \Gamma_c \eta_{t+1} \tag{5.3}$$

$$\Gamma_c = \begin{bmatrix} 1 & 0 & \kappa_c \frac{1 - 1/\psi}{1-\kappa_c \rho} & 0 & 0 & 0 \end{bmatrix} \tag{5.4}$$

$$r_{d,t+1} = \overline{r_d} + \frac{1}{\psi} z_{t+1} + \Gamma_d \eta_{t+1} \tag{5.5}$$

$$\Gamma_d = \begin{bmatrix} 0 & 0 & \kappa_d \frac{\lambda - 1/\psi}{1-\kappa_d \rho} & 0 & 1 & 0 \end{bmatrix} \tag{5.6}$$

$$E_t[r_{c,t+1}^{ex}] = -\Gamma_m S \Gamma_c^{'} - \frac{1}{2} \Gamma_c S \Gamma_c^{'} \tag{5.7}$$

$$E_t[r_{d,t+1}^{ex}] = -\Gamma_m S \Gamma_d^{'} - \frac{1}{2} \Gamma_d S \Gamma_d^{'} \tag{5.8}$$

where $\eta_t = (\sigma_{c,h} \cdot \varepsilon_{c,t}, \sigma_{c,f} \cdot \varepsilon_{c,t}^*, \sigma_{z,h} \cdot \varepsilon_{z,t}, \sigma_{z,f} \cdot \varepsilon_{z,t}^*, \sigma_{d,h} \cdot \varepsilon_{d,t}, \sigma_{d,f} \cdot \varepsilon_{d,t}^*)'$ is the vector of the shocks.

5.2 Limited Information-Learning from Consumption

As derived in section 4.1, one can show the one-step-ahead evolution equations for the variances and covariances of the filtering errors are:

$$Phh_t^c = \frac{\left(\rho^2 {\sigma_{c,h}}^2 \left(Phh_{t-1}^c \left(-{\rho_c}^2 {\sigma_{c,f}}^2 + Pff_{t-1}^c + {\sigma_{c,f}}^2\right) - {Phf_{t-1}^c}^2\right)\right)}{\left({\sigma_{c,f}}^2 \left(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^c + {\sigma_{c,h}}^2\right) - 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} + Pff_{t-1}^c \left(Phh_{t-1}^c + {\sigma_{c,h}}^2\right) - {Phf_{t-1}^c}^2\right)} + {\sigma_{z,h}}^2$$

$$Phf_t^c = \frac{\left(\rho^2 \sigma_{c,h}\sigma_{c,f} \left(\left({\rho_c}^2 - 1\right) Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} + \rho_c \left({Phf_{t-1}^c}^2 - Pff_{t-1}^c Phh_{t-1}^c\right)\right)\right)}{\left(-{\sigma_{c,f}}^2 \left(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^c + {\sigma_{c,h}}^2\right) + 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} - Pff_{t-1}^c \left(Phh_{t-1}^c + {\sigma_{c,h}}^2\right) + {Phf_{t-1}^c}^2\right)} + \rho_z \sigma_{z,f}\sigma_{z,h}$$

$$Pff_t^c = \frac{\left(\rho^2 {\sigma_{c,f}}^2 \left(\left({\rho_c}^2 - 1\right) Pff_{t-1}^c {\sigma_{c,h}}^2 - Pff_{t-1}^c Phh_{t-1}^c + {Phf_{t-1}^c}^2\right)\right)}{\left(-{\sigma_{c,f}}^2 \left(-{\rho_c}^2 {\sigma_{c,h}}^2 + Phh_{t-1}^c + {\sigma_{c,h}}^2\right) + 2\rho_c Phf_{t-1}^c \sigma_{c,h}\sigma_{c,f} - Pff_{t-1}^c \left(Phh_{t-1}^c + {\sigma_{c,h}}^2\right) + {Phf_{t-1}^c}^2\right)} + {\sigma_{z,f}}^2$$

As shown in section 4.1, one can derive the 2 by 2 Kalman gains in this case:

$$K^c = \begin{bmatrix} K_{11,t}^c & K_{12,t}^c \\ K_{21,t}^c & K_{22,t}^c \end{bmatrix}$$

where

$$K_{11,t}^c = \frac{\left(Phf_{t-1}^c(-Phf_{t-1}^c - \rho_c\sigma_{c,h}\sigma_{c,f})\right) + \left(Phh_{t-1}^c \left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right)\right)}{\left(\left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right) - (Phf_{t-1}^c + \rho_c\sigma_{c,h}\sigma_{c,f})^2\right)} \tag{5.9}$$

$$K_{12,t}^c = \frac{\left(Phf_{t-1}^c \left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\right) + \left(Phh_{t-1}^c(-Phf_{t-1}^c - \rho_c\sigma_{c,h}\sigma_{c,f})\right)}{\left(\left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right) - (Phf_{t-1}^c + \rho_c\sigma_{c,h}\sigma_{c,f})^2\right)} \tag{5.10}$$

$$K_{21,t}^c = \frac{\left(Pff_{t-1}^c(-Phf_{t-1}^c - \rho_c\sigma_{c,h}\sigma_{c,f})\right) + \left(Phf_{t-1}^c \left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right)\right)}{\left(\left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right) - (Phf_{t-1}^c + \rho_c\sigma_{c,h}\sigma_{c,f})^2\right)} \tag{5.11}$$

$$K_{22,t}^c = \frac{\left(Pff_{t-1}^c \left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\right) + \left(Phf_{t-1}^c(-Phf_{t-1}^c - \rho_c\sigma_{c,h}\sigma_{c,f})\right)}{\left(\left({\sigma_{c,h}}^2 + Phh_{t-1}^c\right)\left({\sigma_{c,f}}^2 + Pff_{t-1}^c\right) - (Phf_{t-1}^c + \rho_c\sigma_{c,h}\sigma_{c,f})^2\right)} \tag{5.12}$$

the one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components have the following expressions:

$$\widehat{z_t} = \rho \cdot \widehat{z_{t-1}} + K^c_{11,t} \cdot (\sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \widehat{z_t}) + K^c_{12,t} \cdot (\sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \widehat{z^*_t}) \tag{5.13}$$

$$\widehat{z^*_t} = \rho \cdot \widehat{z^*_{t-1}} + K^c_{21,t} \cdot (\sigma_{c,h} \cdot \varepsilon_{c,t} + z_t - \widehat{z_t}) + K^c_{22,t} \cdot (\sigma_{c,f} \cdot \varepsilon^*_{c,t} + z^*_t - \widehat{z^*_t}) \tag{5.14}$$

The Kalman Gains in equation 5.13-5.14 can be steady-state Kalman Gain calculated by solving the Riccati Equation 4.18, or they could be off-steady-state in which case the Kalman Gains themselves are "state" variables, evolving according to equations 5.9-5.12from initial conditions.

Under limited information, the Euler equations 3.2 yield the policy function for price-consumption and price-dividend ratio as a function of the filtered long-run persistent components $\widehat{z_t}, \widehat{z^*_t}$. Utilizing the mapping machinery in section 4.4, one could show that once we define $\eta^c_t = \begin{pmatrix} \nu_{c,t} \\ \nu^*_{c,t} \\ K^c_{11,t}\nu_{c,t} + K^c_{12,t}\nu^*_{c,t} \\ K^c_{21,t}\nu_{c,t} + K^c_{22,t}\nu^*_{c,t} \\ \lambda \cdot \nu_{c,t} \\ \lambda \cdot \nu^*_{c,t} \end{pmatrix}$ and $S_{LC} = \Sigma_{LC} P_{LC} \Sigma'_{LC}$ (4.66), we have the following asset pricing results:

$$
\begin{aligned}
m_{t+1} &= \overline{m} - \frac{1}{\psi}\widehat{z_{t+1}} - \gamma\nu_{c,t+1} - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{11,t+1} + K^c_{12,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c} \\
r_{c,t+1} &= \overline{r_c} + \frac{1}{\psi}\widehat{z_{t+1}} + \nu_{c,t+1} + \frac{\left(1 - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{11,t+1} + K^c_{12,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c} \\
r_{d,t+1} &= \overline{r_d} + \frac{1}{\psi}\widehat{z_{t+1}} + \lambda\nu_{c,t+1} - \frac{\left(\lambda - \frac{1}{\psi}\right)\kappa_d\left(\nu_{c,t+1}K^c_{11,t+1} + K^c_{12,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_d} \\
E_t[r^{ex}_{c,t+1}] &= -\Gamma_m S_{LC}\Gamma_c^{'} - \frac{1}{2}\Gamma_c S_{LC}\Gamma_c^{'} \\
E_t[r^{ex}_{d,t+1}] &= -\Gamma_m S_{LC}\Gamma_d^{'} - \frac{1}{2}\Gamma_d S_{LC}\Gamma_d^{'}
\end{aligned}
\tag{5.15}
$$

5.3 Limited Information-Learning from Consumption and Dividend

The expressions for the evolution equation of Kalman Gain and filtering error are quite involved, and are included in section 4.2.

The one-step-ahead state evolution equations for the filtered home and foreign long-run persistent components are:

$$
\widetilde{z}_t = \rho \cdot \widetilde{z_{t-1}} + K^d_{11,t}\cdot\nu_{c,t} + K^d_{12,t}\cdot\nu^*_{c,t} + K^d_{13,t}\cdot\nu_{d,t} + K^d_{14,t}\cdot\nu^*_{d,t} \tag{5.16}
$$

$$
\widetilde{z}^*_t = \rho \cdot \widetilde{z^*_{t-1}} + K^d_{21,t}\cdot\nu_{c,t} + K^d_{22,t}\cdot\nu^*_{c,t} + K^d_{23,t}\cdot\nu_{d,t} + K^d_{24,t}\cdot\nu^*_{d,t} \tag{5.17}
$$

The Kalman Gains in equation 5.16-5.17 can be steady-state Kalman Gain calculated by solving the Riccati Equation 4.27, or they could be off-steady-state and evolve from initial conditions according to the dynamics derived in section 4.2.

Under limited information, the Euler equations 3.2 yields the policy function for price-consumption and price-dividend ratio as a function of the filtered long-run persistent components $\widehat{z}_t, \widehat{z^*_t}$. Utilizing the mapping machinery in section 4.4, one could show that once we define

$$\eta_t^d = \begin{pmatrix} \nu_{c,t} \\ \nu^*{}_{c,t} \\ K^d_{11,t}\nu_{c,t} + K^d_{13,t}\nu_{d,t} + K^d_{12,t}\nu^*{}_{c,t} + K^d_{14,t}\nu^*{}_{d,t} \\ K^d_{21,t}\nu_{c,t} + K^d_{23,t}\nu_{d,t} + K^d_{22,t}\nu^*{}_{c,t} + K^d_{24,t}\nu^*{}_{d,t} \\ \nu_{c,t} + \nu_{d,t} \\ \nu^*{}_{c,t} + \nu^*{}_{d,t} \end{pmatrix}$$

and $S_{LCD} = \Sigma_{LCD} P_{LCD} \Sigma'_{LCD}$ (4.69),

we have the following asset pricing results:

$$\begin{aligned} m_{t+1} = \overline{m} - \frac{1}{\psi}\widetilde{z_{t+1}} - \gamma\nu_{c,t+1} \\ - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(K^d_{11,t+1}\nu_{c,t+1} + K^d_{13,t+1}\nu_{d,t+1} + K^d_{12,t+1}\nu^*{}_{c,t+1} + K^d_{14,t+1}\nu^*{}_{d,t+1}\right)}{1-\rho\kappa_c} \end{aligned} \tag{5.18}$$

$$\begin{aligned} r_{c,t+1} = \overline{r_c} + \frac{1}{\psi}\widetilde{z_{t+1}} + \nu_{c,t+1} \\ + \frac{\left(1 - \frac{1}{\psi}\right)\kappa_c\left(K^d_{11,t+1}\nu_{c,t+1} + K^d_{13,t+1}\nu_{d,t+1} + K^d_{12,t+1}\nu^*{}_{c,t+1} + K^d_{14,t+1}\nu^*{}_{d,t+1}\right)}{1-\rho\kappa_c} \end{aligned} \tag{5.19}$$

$$\begin{aligned} r_{d,t+1} = \overline{r_d} + \frac{1}{\psi}\widetilde{z_{t+1}} + \nu_{c,t+1} + \nu_{d,t+1} \\ - \frac{\left(\lambda - \frac{1}{\psi}\right)\kappa_d\left(\nu_{c,t+1}K^d_{11,t+1} + K^d_{13,t+1}\nu_{d,t+1} + K^d_{12,t+1}\nu^*{}_{c,t+1} + K^d_{14,t+1}\nu^*{}_{d,t+1}\right)}{1-\rho\kappa_d} \end{aligned} \tag{5.20}$$

and the conditional expectations of excess returns are given by

$$E_t[r^{ex}_{c,t+1}] = -\Gamma_m S_{LCD}\Gamma'_c - \frac{1}{2}\Gamma_c S_{LCD}\Gamma'_c$$
$$E_t[r^{ex}_{d,t+1}] = -\Gamma_m S_{LCD}\Gamma'_d - \frac{1}{2}\Gamma_d S_{LCD}\Gamma'_d$$

Chapter 6

Results

6.1 Contagion

In this section I will present both theoretical and simulation results. By utilizing the mapping machinery in section 4.4 and the log-linearization results developed in section 4.3, one can derive the following theoretical expressions for the pricing kernel as in section 5,

For the information structure of learning from consumption stream, we have:

$$
\begin{aligned}
m_{t+1} &= \overline{m} - \frac{1}{\psi}\widehat{z_{t+1}} - \gamma\nu_{c,t+1} - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{11,t+1} + K^c_{12,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c} \\
m^*_{t+1} &= \overline{m}^* - \frac{1}{\psi}\widehat{z_{t+1}}^* - \gamma\nu^*_{c,t+1} - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{21,t+1} + K^c_{22,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c}
\end{aligned}
\tag{6.1}
$$

For the information structure of learning from consumption and dividend stream, we have:

$$
\begin{aligned}
m_{t+1} &= \overline{m} - \frac{1}{\psi}\widetilde{z_{t+1}} - \gamma\nu_{c,t+1} \\
&\quad - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^d_{11,t+1} + K^d_{13,t+1}\nu_{d,t+1} + K^d_{12,t+1}{\nu^*}_{c,t+1} + K^d_{14,t+1}{\nu^*}_{d,t+1}\right)}{1-\rho\kappa_c}
\end{aligned}
\tag{6.2}
$$

$$
\begin{aligned}
m^*_{t+1} &= \overline{m}^* - \frac{1}{\psi}\widetilde{z_{t+1}}^* - \gamma\nu^*_{c,t+1} \\
&\quad - \frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^d_{21,t+1} + K^d_{23,t+1}\nu_{d,t+1} + K^d_{22,t+1}{\nu^*}_{c,t+1} + K^d_{24,t+1}{\nu^*}_{d,t+1}\right)}{1-\rho\kappa_c}
\end{aligned}
$$

There is no unique definition in international asset pricing literature as to what contagion is. It can be viewed as foreign influence on US equity risk premium (Chan, Karolyi, and Stulz, 1992). It can be viewed as the documented phenomenon that shocks to Nikkei Index can impact S&P 500 and vice-versa (Karolyi and Stulz, 1996). It can be defined as how regional markets respond to information in one country, such as the crisis Asia in 1997 and Latin American in 1994 (Bekaert and Harvey, 2003). However, Bekaert, Harvey, and Ng (2005) maintains that contagion should be more than just what is revealed as increased correlation during crisis, since statistically it is natural for correlation to go up when volatility is high (Ang and Bekaert, 2002). The preferred definition of contagion is the propagation of shocks in excess of what can be explained by fundamentals (Forbes and Rigobon, 2000). This is the definition of contagion I use in this article, and I define economic fundamentals as the long-run persistent components.

Under full information, the pricing kernel of home only depends on home's long-run component. Under limited information with learning, one could readily see how contagion occurs by observing equation 6.1 and equation 6.2. The home pricing kernel will respond to foreign innovation in consumption $\nu^*_{c,t+1}$, and when learning from dividend stream as well, home pricing kernel will also respond to foreign innovation in dividend $\nu^*_{d,t+1}$. Similarly, foreign pricing kernel will respond to home innovations. Since markets are complete and there is no arbitrage, all assets are priced using the pricing kernel. Thus home asset price will drop in response to foreign bad news and rise in response to foreign good news, even when there is no shock to home fundamental.

From simulation results, the Impulse Response Functions (IRF) plots can be used to illustrate this point. Figure 1 shows that under full information, home pricing kernel, and associated asset prices such as risk premium and risk free rate does not respond at all to foreign shock. However, when there is learning involved, both figure 2 and figure 3 show that either short-run or long-run or dividend shock from foreign country can impact home's pricing kernel and returns for an extended period of time.

From the point of view of an agent who uses Kalman filtering to Bayesian update the fundamentals, when innovation in consumption stream occurs, the agent does not know the current economic fundamental and cannot break down the innovation into a short-run component which would affect the economy for only one-period and a long-run component which would affect the economy for many periods. The agent can only optimize asset pricing decisions using her belief of the distribution of the fundamental, formed through using all historical information up to that point in time. However, from the point of view of an econometrician who observe the entire data series and estimate the underlying economic fundamentals at each point in time, she is in a better position to tell if the shock was short term or long term. In other words, there is arbitrage to eyes of the econometrician but not to the agents, since the information set of the econometrician who observe the entire history is larger than the information set of the rational agents who at each point in time observe information only up to that point.

6.2 Forward Premium Puzzle

For compactness of presentation, I will focus on the information structure of learning from consumption stream; all the results also follow for the case of learning from consumption and dividend stream. To solve the Forward Premium Puzzle, we need to establish the relationship between P^c_{hh} and the off-steady-state Kalman Gains. It can be shown in the comparative statics results that $\frac{\partial K^c_{11,t}}{\partial P^c_{hh}}$ is positive and orders of magnitude larger than $\frac{\partial K^c_{12,t}}{\partial P^c_{hh,t-1}}$, $\frac{\partial K^c_{21,t}}{\partial P^c_{hh,t-1}}$, $\frac{\partial K^c_{22,t}}{\partial P^c_{hh,t-1}}$. This is intuitive because off steady-state, $P^c_{hh,t-1}$ is the agent's belief of the variance of the filtering error of $\widehat{z_{t-1}}$. $P^c_{hh,t-1}$ has the largest effect on $K^c_{11,t}$ which measures how the agent update her estimate of home long-run persistent component in response to home consumption innovation. The intuition behind $\frac{\partial K^c_{11,t}}{\partial P^c_{hh,t-1}} > 0$ is that if $P^c_{hh,t-1}$ is high, then the agent doesn't trust her previous estimate $\widehat{z_{t-1}}$ as much, thus there is a need to revise it based on new innovation, and the Kalman Gain $K^c_{11,t}$ is high for next period; conversely, if $P^c_{hh,t-1}$ is low, then the agent is already quite confident with her estimate $\widehat{z_{t-1}}$ which is based on all historical information, and therefore there is little need to revise it, thus $K^c_{11,t}$ is low.

Suppose the agents of both countries start off at an off steady-state prior regarding the filtering error variance of home country's, but not foreign country's long-run persistent component. Let $P^c_{hh,t_0} = 150\% \times P^c_{hh,ss}$, where $P^c_{hh,ss}$ is the steady state prior obtained from Riccati equation 4.18. Since the agents are learning rationally, through Kalman filter the agents will find out this prior was set to be too high. The consequence of the Kalman filter dynamics is that $lim_{t->\infty} P^c_{hh,t} = P^c_{hh,ss}$. In other words, $P^c_{hh,t}$ is decreasing over time on a deterministic path. We can see from how this generates endogenous time-varying volatility in the pricing kernel from equation 6.1. From section 5.2, we also have the following analytical log-linearized asset pricing equations:

$$r_{c,t+1} = \overline{r_c} + \frac{1}{\psi}\widehat{z_{t+1}} + \nu_{c,t+1} + \frac{\left(1-\frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{11,t+1} + K^c_{12,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c} \tag{6.3}$$

$$r^*_{c,t+1} = \overline{r_c}^* + \frac{1}{\psi}\widehat{z_{t+1}}^* + \nu^*_{c,t+1} + \frac{\left(1-\frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}K^c_{21,t+1} + K^c_{22,t+1}\nu^*_{c,t+1}\right)}{1-\rho\kappa_c} \tag{6.4}$$

The expected change of next period's exchange rate growth is : $E_t[\Delta e_{t+1}] = E_t[m^*_{t+1} - m_{t+1}]$, which by properties of log-normal distribution and equation 6.1 is:

$$\begin{aligned}
E_t[\Delta e_{t+1}] = \overline{m}^* - \overline{m} + E_t \Bigg[& \frac{1}{\psi}\left(\widehat{z_{t+1}}^* - \widehat{z_{t+1}}\right) \\
& - \frac{\left(\gamma-\frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}\left[K^c_{21,t+1} - K^c_{11,t+1}\right] + \nu^*_{c,t+1}\left[K^c_{22,t+1} - K^c_{12,t+1}\right]\right)}{1-\rho\kappa_c}\Bigg] \\
+ \frac{1}{2}V_t \Bigg[& \frac{1}{\psi}\left(\widehat{z_{t+1}}^* - \widehat{z_{t+1}}\right) \\
& - \frac{\left(\gamma-\frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}\left[K^c_{21,t+1} - K^c_{11,t+1}\right] + \nu^*_{c,t+1}\left[K^c_{22,t+1} - K^c_{12,t+1}\right]\right)}{1-\rho\kappa_c}\Bigg]
\end{aligned} \tag{6.5}$$

$$\begin{aligned}
= \overline{m}^* - \overline{m} + \frac{1}{\psi}E_t\left[\left(\widehat{z_{t+1}}^* - \widehat{z_{t+1}}\right)\right] + \frac{1}{2}V_t \Bigg[& \frac{1}{\psi}\left(\widehat{z_{t+1}}^* - \widehat{z_{t+1}}\right) \\
& - \frac{\left(\gamma-\frac{1}{\psi}\right)\kappa_c\left(\nu_{c,t+1}\left[K^c_{21,t+1} - K^c_{11,t+1}\right] + \nu^*_{c,t+1}\left[K^c_{22,t+1} - K^c_{12,t+1}\right]\right)}{1-\rho\kappa_c}\Bigg]
\end{aligned} \tag{6.6}$$

I include the second order terms as well for the forward premium $f_t - s_t = r_{f,t} - r^*_{f,t}$ (Backus, Foresi, and Telmer, 2002).

$$
\begin{aligned}
& r_{f,t+1} - r^*_{f,t+1} && (6.7)\\
&= \overline{r_f} - \overline{r_f}^* + \frac{1}{\psi} E_t\left(\widehat{z_{t+1}} - \widehat{z_{t+1}}^*\right) + \frac{1}{2\psi^2}\left(V_t[\widehat{z_{t+1}}^*] - V_t[\widehat{z_{t+1}}]\right) && (6.8)\\
&= \overline{r_f} - \overline{r_f}^* + \frac{1}{\psi} E_t\left(\widehat{z_{t+1}} - \widehat{z_{t+1}}^*\right) + \frac{1}{2\psi^2}\left(P^c_{ff,t} - P^c_{hh,t}\right) && (6.9)
\end{aligned}
$$

The β_{UIP} coefficient is obtained from the following regression: $\Delta e_{t+1} = \alpha + \beta_{UIP} \cdot (r_{f,t} - r^*_{f,t}) + \varepsilon_t$. Notice the $\frac{1}{\psi} E_t\left(\widehat{z_{t+1}} - \widehat{z_{t+1}}^*\right)$ term which appears in both $r_{f,t+1} - r^*_{f,t+1}$ and $E_t[\Delta e_{t+1}]$; one can readily see that without time-varying volatility, the β_{UIP} coefficient is equal to one (Backus, Foresi, and Telmer, 2002).

Since the comparative statics results show the predominant effect of $P^c_{hh,t-1}$ is on $K^c_{11,t}$, I will group the nuisance terms as Y_{t+1} and simplify the above expression as:

$$
E_t[\Delta e_{t+1}] = \overline{m}^* - \overline{m} + \frac{1}{\psi} E_t\left(\widehat{z_{t+1}} - \widehat{z_{t+1}}^*\right) + V_t\left[-\frac{\left(\gamma - \frac{1}{\psi}\right)\kappa_c\left(-\nu_{c,t+1} K^c_{11,t+1}\right)}{1 - \rho\kappa_c} + Y_{t+1}\right]
$$

Thus we can see as time goes by, $P^c_{hh,t-1}$ is decreasing, $K^c_{11,t}$ is also decreasing, and therefore$E_t[\Delta e_{t+1}]$ is decreasing. The intuition is that as the estimate $\widehat{z}_t$ gets more reliable, there is less risk in home, and home's currency appreciates. For the interest rate differential however, since $P^c_{hh,t}$ is decreasing, $r_{f,t+1} - r^*_{f,t+1}$ is increasing. The intuition is that as the estimate $\widehat{z}_t$ gets more reliable, home has less incentive to invest and demands higher risk free rate.

Thus the above dynamics generates negative β_{UIP} coefficient. This provides a solution to the forward premium puzzle.

6.3 Numerical Results

I performed third order numerical approximations for 25 chains of simulations, each with 700 periods. The parameters used in the calibrations are listed in table 1. Variables of interest have the following symbols: z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

As shown in table 10, learning can elevate risk premium since under learning, the pricing kernel is more volatile due to uncertainty regarding the underlying latent variables. Table 11 shows that the pricing kernel is less volatile under learning from both consumption and dividend stream, compared to learning from consumption stream only, since learning from multiple information sources increases the accuracy of estimated latent variables, which is documented in the filtering error covariance matrices, table 8 and table 9. It is notable that the exchange rate volatility does not go up much compared to the full-information case, since learning generates higher cross-country correlation of the estimated long-run components. In the data between US and UK from 1979 to 2006, the volatility of exchange rate is around 11% (Colacito and Croce, 2011).

Another consequence of learning is that the cross-country correlation of risk free rate $r_{f,t}$, and return on the asset which pays consumption stream $r_{c,t}$, are now higher, as shown in table 13. This effect is more prominent for the case of learning from consumption, and less prominent for the case of learning from consumption and dividend. This is also due to that learning from multiple sources of information enables the agents to better infer the latent variables with accuracy and thus move closer toward full information case.

The fact that agents cannot distinguish short-run from long-run shock can also be observed in the correlation matrix of mapped shocks in table 4 and table 5, which were obtained through the mapping procedure developed in section 4.4. While the true correlation matrix under full information (table 3) has zero correlation between different types of shocks, that is no longer true under learning.

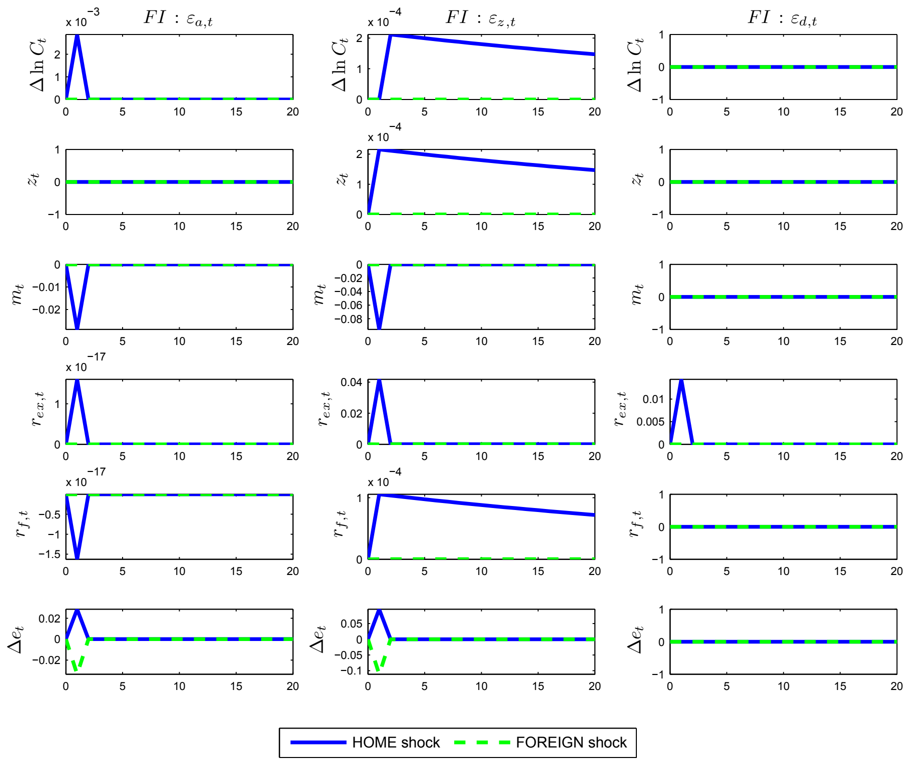
Chapter 7

Concluding Remarks

This article documents the asset pricing implications of learning in consumption-based long-run risk literature in the context of international markets. There are two countries and the long-run prospects of them are not directly observable, though they are known to be correlated. Agents recursively Bayesian update their estimates of long-run persistent components through Kalman filter. Learning provides an explanation for the contagion phenomenon, defined as changes in one country's asset prices in response to foreign news, that occurs in the absence of domestic news. Learning generates higher correlation of pricing kernels across two countries, and higher risk premium without increasing volatility of exchange rate. When evaluated off steady-state, learning can generate time-varying volatility in risk premium and can also provide an explanation for the forward premium puzzle.

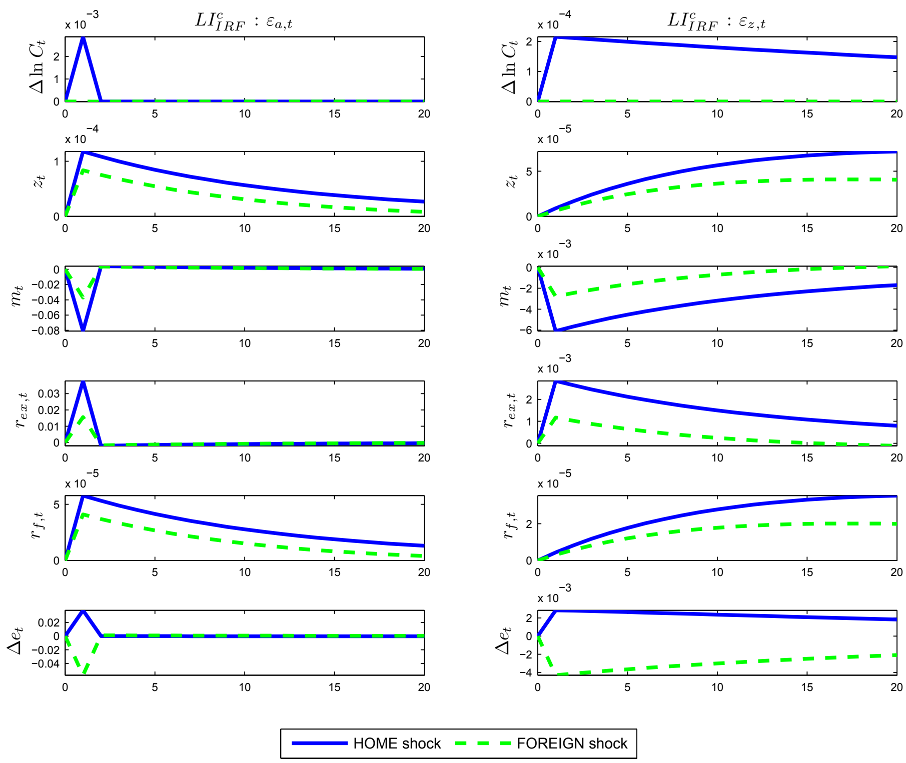
It has been observed in data that correlation across countries is higher in bear markets than in bull markets (Ang and Bekaert, 2002). For future research, one possible extension is to incorporate pessimistic beliefs in the learning process, which will generate even richer asset pricing dynamics. It may be also interesting to extend the model to three countries.

Figure 1: Impulse responses implied in the information structure of Full Information



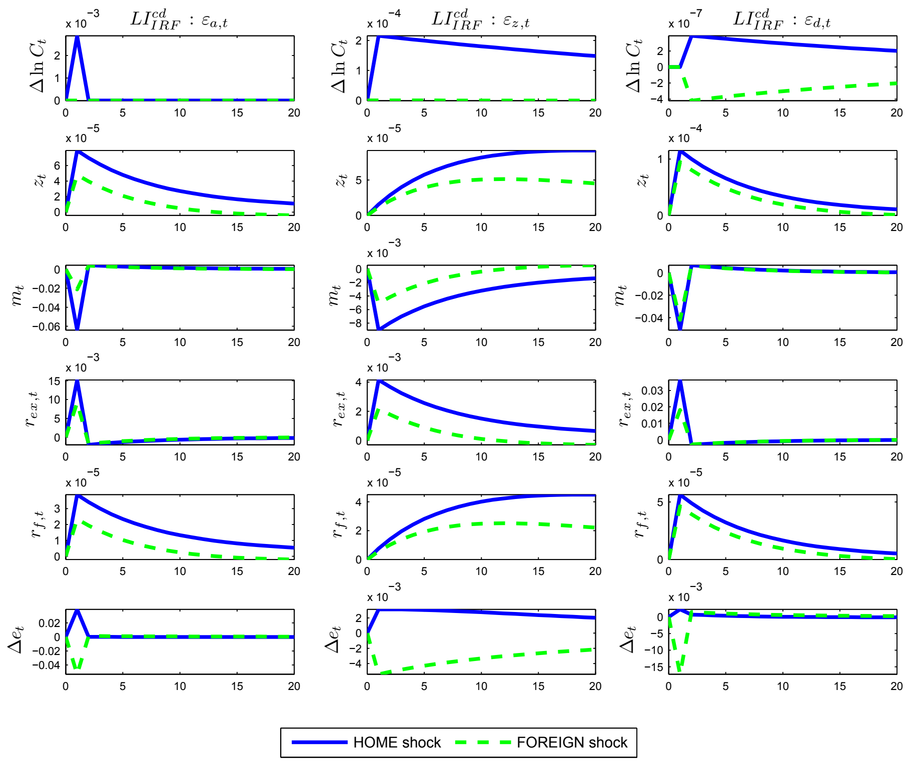
Each shock is given one standard deviation of positive impulse. $\varepsilon_{a,t}$ is short-run shock, $\varepsilon_{z,t}$ is long-run shock, $\varepsilon_{d,t}$ is dividend shock. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

Figure 2: Impulse responses implied in the information structure of Learning from Consumption



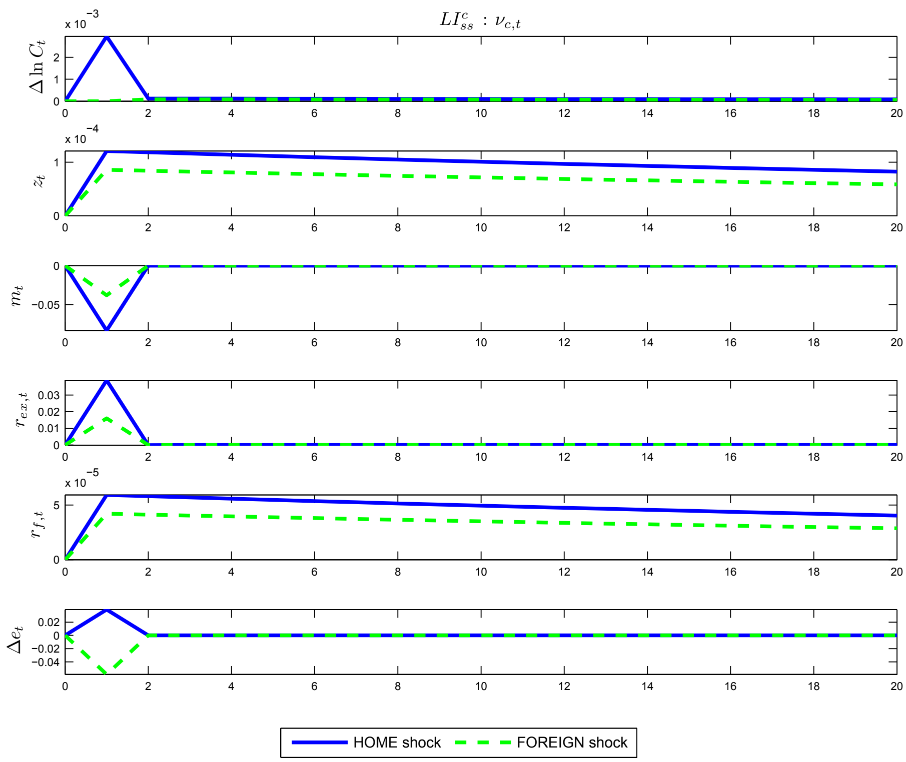
Each shock is given one standard deviation of positive impulse. $\varepsilon_{a,t}$ is short-run shock, $\varepsilon_{z,t}$ is long-run shock. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

Figure 3: Impulse responses implied in the information structure of Learning from Consumption and Dividend



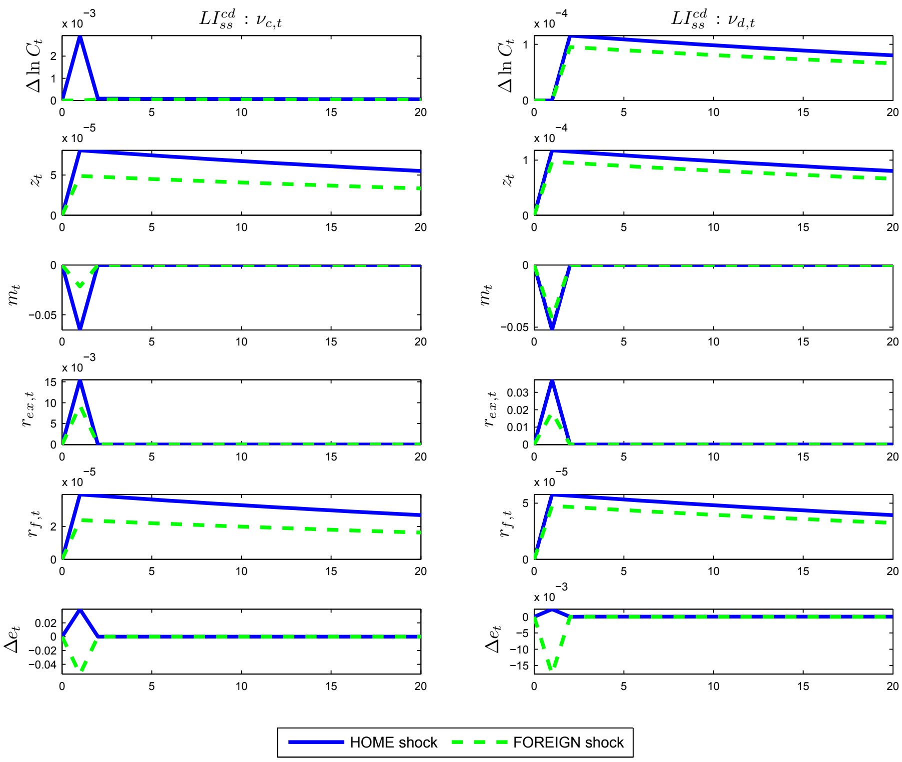
Each shock is given one standard deviation of positive impulse. $\varepsilon_{a,t}$ is short-run shock, $\varepsilon_{z,t}$ is long-run shock, $\varepsilon_{d,t}$ is dividend shock. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

Figure 4: Impulse responses to innovation shocks implied in the information structure of Learning from Consumption



Each shock is given one standard deviation of positive impulse. $\nu_{c,t}$ is consumption innovation shock. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

Figure 5: Impulse responses to innovation shocks implied in the information structure of Learning from Consumption and Dividend



Each shock is given one standard deviation of positive impulse. $\nu_{c,t}$ is consumption innovation shock, $\nu_{d,t}$ is dividend innovation shock. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream.

Table 1: Parameters used for Calibration

Parameter	Value
Persistence of long run shock ρ	0.98
Subjective discount factor δ	0.9985
Long run mean of Consumption Growth μ	0.0015
Long run mean of Dividend Growth μ_d	0.0007
Risk aversion γ	10
Intertemporal elasticity of substituion ψ	2
Leverage. Dividend process LongRun component multiple λ	5.5
Composite Parameter θ	-0.055556
Ratio of Dividend shock and Short-Run shock volatilities φ_d	5
Correlation of Home/Foregin Short-Run shock $\rho\left(\varepsilon_{a,t}, \varepsilon^*_{a,t}\right)$	0.35
Correlation of Home/Foregin Long-Run shock $\rho\left(\varepsilon_{z,t}, \varepsilon^*_{z,t}\right)$	0.93
Correlation of Home/Foregin Dividend shock $\rho\left(\varepsilon_{d,t}, \varepsilon^*_{d,t}\right)$	-0.1
Variance of Home Short-Run shock $\sigma^2_{a,t}$	8.28e-006
Covariance of Home/Foreign Short-Run shock $\sigma_{a,a^*,t}$	3.381e-006
Variance of Foregin Short-Run shock $\sigma^2_{a^*,t}$	1.127e-005
Variance of Home Long-Run shock $\sigma^2_{z,t}$	4.6656e-008
Covariance of Home/Foreign Long-Run shock $\sigma_{z,z^*,t}$	5.0622e-008
Variance of Foregin Long-Run shock $\sigma^2_{z^*,t}$	6.3504e-008
Variance of Home Dividend shock $\sigma^2_{d,t}$	0.000207
Covariance of Home/Foregin Dividend shock $\sigma_{d,d^*,t}$	-2.415e-005
Variance of Foregin Dividend shock $\sigma^2_{d^*,t}$	0.00028175

Table 2: Variance-Covariance Matrix under Full Information

	$\varepsilon_{a,t}$	$\varepsilon^*_{a,t}$	$\varepsilon_{z,t}$	$\varepsilon^*_{z,t}$	$\varepsilon_{d,t}$	$\varepsilon^*_{d,t}$
$\varepsilon_{a,t}$	8.28e-006	3.381e-006	0	0	0	0
$\varepsilon^*_{a,t}$	3.381e-006	1.127e-005	0	0	0	0
$\varepsilon_{z,t}$	0	0	4.6656e-008	5.0622e-008	0	0
$\varepsilon^*_{z,t}$	0	0	5.0622e-008	6.3504e-008	0	0
$\varepsilon_{d,t}$	0	0	0	0	0.000207	-2.415e-005
$\varepsilon^*_{d,t}$	0	0	0	0	-2.415e-005	0.00028175

Table 3: True Correlation Matrix under Full Information

	$\varepsilon_{a,t}$	$\varepsilon^*_{a,t}$	$\varepsilon_{z,t}$	$\varepsilon^*_{z,t}$	$\varepsilon_{d,t}$	$\varepsilon^*_{d,t}$
$\varepsilon_{a,t}$	1	0.35	0	0	0	0
$\varepsilon^*_{a,t}$	0.35	1	0	0	0	0
$\varepsilon_{z,t}$	0	0	1	0.93	0	0
$\varepsilon^*_{z,t}$	0	0	0.93	1	0	0
$\varepsilon_{d,t}$	0	0	0	0	1	-0.1
$\varepsilon^*_{d,t}$	0	0	0	0	-0.1	1

Table 4: Correlation Matrix for Mapped Shocks under Learning from Consumption

	$\hat{\varepsilon}_{a,t}$	$\hat{\varepsilon}^*_{a,t}$	$\hat{\varepsilon}_{z,t}$	$\hat{\varepsilon}^*_{z,t}$
$\hat{\varepsilon}_{a,t}$	1	0.37639	0.88709	0.76154
$\hat{\varepsilon}^*_{a,t}$	0.37639	1	0.76154	0.88709
$\hat{\varepsilon}_{z,t}$	0.88709	0.76154	1	0.97473
$\hat{\varepsilon}^*_{z,t}$	0.76154	0.88709	0.97473	1

Table 5: Correlation Matrix for Mapped Shocks under Learning from Consumption and Dividend

	$\hat{\varepsilon}_{a,t}$	$\hat{\varepsilon}^*_{a,t}$	$\hat{\varepsilon}_{z,t}$	$\hat{\varepsilon}^*_{z,t}$	$\hat{\varepsilon}_{d,t}$	$\hat{\varepsilon}^*_{d,t}$
$\hat{\varepsilon}_{a,t}$	1	0.36687	0.56191	0.45826	0.039716	0.03239
$\hat{\varepsilon}^*_{a,t}$	0.36687	1	0.45826	0.56191	0.03239	0.039716
$\hat{\varepsilon}_{z,t}$	0.56191	0.45826	1	0.97112	0.61576	0.50218
$\hat{\varepsilon}^*_{z,t}$	0.45826	0.56191	0.97112	1	0.50218	0.61576
$\hat{\varepsilon}_{d,t}$	0.039716	0.03239	0.61576	0.50218	1	-0.060154
$\hat{\varepsilon}^*_{d,t}$	0.03239	0.039716	0.50218	0.61576	-0.060154	1

Table 6: Numerical value of steady-state Kalman Gain Matrix under Learning from Consumption

0.040927	0.024985
0.034007	0.040927

Table 7: Numerical value of steady-state Kalman Gain Matrix under Learning from Consumption and Dividend

0.027375	0.014482	0.0079858	0.0056773
0.019711	0.027375	0.0077274	0.0079858

Table 8: Numerical value of steady-state Covariance matrix of the filtering errors under Learning from Consumption

	$\widehat{\varepsilon}_{h,t}$	$\widehat{\varepsilon}_{f,t}$
$\widehat{\varepsilon}_{h,t}$	4.5324e-007	4.5394e-007
$\widehat{\varepsilon}_{f,t}$	4.5394e-007	6.1691e-007

Table 9: Numerical value of steady-state Covariance matrix of the filtering errors under Learning from Consumption and Dividend

	$\widetilde{\varepsilon}_{h,t}$	$\widetilde{\varepsilon}_{f,t}$
$\widetilde{\varepsilon}_{h,t}$	3.1137e-007	2.9626e-007
$\widetilde{\varepsilon}_{f,t}$	2.9626e-007	4.2381e-007

Table 10: Numerical Simulation Results: Annualized Mean

	FI	LI_{ss}^{c}	LI_{ss}^{cd}
z_t	0	0	0
$\Delta \ln C_t$	1.8	1.8	1.8
$\Delta \ln D_t$	0.84	0.84	0.84
m_t	-8.4586	-8.762	-8.7359
Δe_t	-2.0553	-2.1795	-2.1791
$r_{f,t}$	2.3328	2.1909	2.231
$r_{d,t}$	6.4177	7.2046	6.6513
$r_{ex,t}$	4.0849	5.0137	4.4204
$\ln \frac{P_t}{C_t}$	79.2296	77.8268	77.9568
$r_{c,t}$	3.0044	3.0205	3.019

FI is full information benchmark; LI_{ss}^{c} is learning from consumption stream using steady-state Kalman filter; LI_{ss}^{cd} is learning from consumption and dividend stream using steady-state Kalman filter. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

Table 11: Numerical Simulation Results: Annualized Volatility

	FI	LI^{c}_{ss}	LI^{cd}_{ss}
z_t	0.3218	0.2695	0.2999
$\Delta \ln C_t$	1.0444	1.0696	1.0589
$\Delta \ln D_t$	5.1658	5.8829	5.2499
m_t	34.85	36.4356	35.6694
Δe_t	19.2524	19.9782	20.2438
$r_{f,t}$	0.1576	0.132	0.1469
$r_{d,t}$	15.4034	16.6043	16.1085
$r_{ex,t}$	15.4068	16.6067	16.1112
$\ln \frac{P_t}{C_t}$	7.5098	6.284	6.9969
$r_{c,t}$	2.012	2.3885	2.2648

FI is full information benchmark; LI^{c}_{ss} is learning from consumption stream using steady-state Kalman filter; LI^{cd}_{ss} is learning from consumption and dividend stream using steady-state Kalman filter. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

Table 12: Numerical Simulation Results: Autocorrelation-ACF(1)

	FI	LI^c_{ss}	LI^{cd}_{ss}
z_t	96.9764	97.0699	97.5444
$\Delta \ln C_t$	10.3912	8.1219	11.2292
$\Delta \ln D_t$	13.1464	8.1219	10.7614
m_t	1.8545	-1.4132	0.7303
Δe_t	-3.7041	0.2235	0.9167
$r_{f,t}$	96.9764	97.0709	97.5447
$r_{d,t}$	4.4295	-1.3378	0.7313
$r_{ex,t}$	4.2478	-1.4892	0.5728
$\ln \frac{P_t}{C_t}$	96.9759	97.07	97.5442
$r_{c,t}$	2.4978	-0.2555	2.1265

FI is full information benchmark; LI^c_{ss} is learning from consumption stream using steady-state Kalman filter; LI^{cd}_{ss} is learning from consumption and dividend stream using steady-state Kalman filter. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

Table 13: Numerical Simulation Results: Correlation of Home/Foreign Counterparts

	FI	LI_{ss}^{c}	LI_{ss}^{cd}
z_t	92.1642	97.6228	96.0112
$\Delta \ln C_t$	39.5981	40.9778	39.2543
$\Delta \ln D_t$	1.9399	40.9778	0.7725
m_t	88.3361	88.0758	87.1885
Δe_t	-100	-100	-100
$r_{f,t}$	92.1937	97.6037	96.0116
$r_{d,t}$	82.1113	84.8666	80.5397
$r_{ex,t}$	82.1154	84.8671	80.5463
$\ln \frac{P_t}{C_t}$	79.262	93.2036	87.8564
$r_{c,t}$	71.2044	75.3846	72.3027

FI is full information benchmark; LI_{ss}^{c} is learning from consumption stream using steady-state Kalman filter; LI_{ss}^{cd} is learning from consumption and dividend stream using steady-state Kalman filter. z_t is the long-run persistent component; ΔlnC_t is log consumption growth; ΔlnD_t is log dividend growth; m is pricing kernel; Δe_t is log exchange rate growth; $r_{f,t}$ is log risk free rate; $r_{d,t}$ is log return on dividend stream; $r_{ex,t}$ is equity risk premium; $ln\frac{P_t}{C_t}$ is the log price consumption ratio ; $r_{c,t}$ is the log return on the asset which pays the consumption stream. Simulation is done with third order numerical approximation for 25 chains each with 700 periods in monthly frequency but reported values are annualized and multiplied by 100.

Table 14: Numerical Simulation Results: Theoretical β_{UIP} Regression Coefficients

	FI	LI_{ss}^{c}	LI_{ss}^{cd}	LI_{hh+}^{c}	LI_{hh+}^{cd}	$Data$
β_{UIP}	1.00	1.00	1.00	-2.84	-0.74	-0.72

Theoretical β_{UIP} regression coefficients of the equation $\Delta e_{t+1} = \alpha + \beta_{UIP} \cdot (r_{f,t} - r^*_{f,t}) + \varepsilon_t$. FI is full information benchmark; LI_{hh+}^{c} is learning from consumption stream using a not yet converged Kalman filter from a prior in which variance of estimation error of home's latent variable was believed to be 150% the steady-state Kalman filter value; LI_{hh+}^{cd} is learning from consumption and dividend stream using a not yet converged Kalman filter from a prior in which variance of estimation error of home's latent variable was believed to be 150% the steady-state Kalman filter value. Simulation is done with third order numerical approximation for 25 chains each with 100 periods in monthly frequency.

Bibliography

Abel, A.B., 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

Ai, Chunrong, and Xiaohong Chen, 2003, Efficient estimation of models with conditional moment restrictions containing unknown functions, *Econometrica* 71, 1795–1843.

——— , 2007, Estimation of possibly misspecified semiparametric conditional moment restriction models with different conditioning variables, *Journal of Econometrics* 141, 5–43.

Ai, H., 2010, Information quality and long-run risk: Asset pricing implications, *The Journal of Finance* 65, 1333–1367.

Ai, Hengjie, Mariano Massimiliano Croce, and Kai Li, 2010, Toward a quantitative general equilibrium asset pricing model with intangible capital, Unpublished paper, UNC Kenan-Flagler Business School.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, 1998, Risk and robustness in equilibrium, Unpublished paper, University of Chicago.

——— , 2003, A quartet of semigroups for model specification, robustness, prices of risk, and model detection, *Journal of the European Economic Association* 1, 68–123.

Ang, A., and G. Bekaert, 2002, International asset allocation with regime shifts, *Review of Financial studies* 15, 1137–1187.

Backus, D.K., S. Foresi, and C.I. Telmer, 2002, Affine term structure models and the forward premium anomaly, *The Journal of Finance* 56, 279–304.

Backus, D.K., B.R. Routledge, and S.E. Zin, 2005, Exotic preferences for macroeconomists, in *NBER Macroeconomics Annual 2004, Volume 19* . pp. 319–414 (MIT Press).

Bansal, R., R.F. Dittmar, and C.T. Lundblad, 2005, Consumption, dividends, and the cross section of equity returns, *The Journal of Finance* 60, 1639–1672.

Bansal, R., D. Kiku, and A. Yaron, 2010, Long run risks, the macroeconomy, and asset prices, in *American Economic Review: Papers & Proceedings* vol. 100 pp. 1343–1375.

Bansal, R., and I. Shaliastovich, 2009, Confidence risk and asset prices, Discussion paper, National Bureau of Economic Research.

——— , 2011, Learning and asset-price jumps, *Review of Financial Studies* 24, 2738–2780.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Bekaert, G., and C.R. Harvey, 2003, Emerging markets finance, *Journal of Empirical Finance* 10, 3–55.

——— , and A. Ng, 2005, Market integration and contagion, *Journal of Business* 78.

Chan, K.C., G.A. Karolyi, and R.M. Stulz, 1992, Global financial markets and the risk premium on us equity, *Journal of Financial Economics* 32, 137–167.

Chen, X., J. Favilukis, and S.C. Ludvigson, 2011, An estimation of economic models with recursive preferences, Discussion paper, National Bureau of Economic Research.

Cipriani, M., and A. Guarino, 2008, Herd behavior and contagion in financial markets, *The BE Journal of Theoretical Economics* 8.

Colacito, Riccardo, and Mariano M. Croce, 2010, The Short- and Long-Run Benefits of Financial Integration, *The American Economic Review* 100(2), 527–31.

——— , 2011, Risks for the Long Run and the Real Exchange Rate, *Journal of Political Economy* 119(1), 153–182.

Croce, M.M., M. Lettau, and S.C. Ludvigson, 2012, Investor information, long-run risk, and the term structure of equity, *Working Paper*.

Epstein, Larry G, and Stanley E Zin, 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937–69.

——— , 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy* 99, 263–86.

Erb, C.B., C.R. Harvey, and T.E. Viskanta, 1998, Contagion and risk, *Emerging Markets Quarterly* 2, 46–64.

Fama, E.F., 1984, Forward and spot exchange rates, *Journal of Monetary Economics* 14, 319–338.

Forbes, K., and R. Rigobon, 2000, Contagion in latin america: Definitions, measurement, and policy implications, Discussion paper, National Bureau of Economic Research.

Fromlet, H., 2001, Behavioral finance-theory and practical application, *Business Economics* 36, 63–69.

Hansen, L.P., and T.J. Sargent, 2010, Fragile beliefs and the price of uncertainty, *Quantitative Economics* 1, 129–162.

Hansen, Lars Peter, 2007, Beliefs, doubts and learning: Valuing macroeconomic risk, *American Economic Association Papers and Proceedings* Richard T. Ely Lecture 97, 1–30.

——— , John Heaton, and Nan Li, 2008, Consumption strikes back?: Measuring long-run risk, *Journal of Political Economy* 116, 260–302.

Harvey, Andrew C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press: Cambridge, United Kingdon).

Haykin, S.S., 2001, *Kalman filtering and neural networks* (Wiley Online Library).

Johannes, M., L. Lochstoer, and Y. Mou, 2010, Learning about consumption dynamics, in *AFA 2011 Denver Meetings Paper*.

Kalman, R.E., 1960, Contributions to the theory of optimal control, *Bol. Soc. Mat. Mex* 5, 102–119.

——— , and R.S. Bucy, 1961, New results in linear filtering and prediction theory, *Journal of Basic Engineering* 83, 95–108.

Karolyi, G.A., and R.M. Stulz, 1996, Why do markets move together? an investigation of us-japan stock return comovements, *The Journal of Finance* 51, 951–986.

Lawler, G.F., 2010, *Random walk and the heat equation* . vol. 55 (American Mathematical Society).

Murphy, K.P., 2002, Dynamic bayesian networks: representation, inference and learning, Ph.D. thesis University of California.

Pastor, L., and P. Veronesi, 2009, Learning in financial markets, Discussion paper, National Bureau of Economic Research.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, 2007, Asset pricing with idiosyncratic risk and overlapping generations, *Review of Economic Dynamics* 10, 519–548.

Telmer, Chris I., 1993, Asset-pricing puzzles and incomplete markets, *Journal of Finance* 48, 1803–1832.

Weil, Philippe, 1989, The equity premium puzzle and the riskfree rate puzzle, *Journal of Monetary Economics* 24, 401–422.